12-27-99

Docket # PF-0519-1 DIV

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
REQUEST FOR FILING A PATENT APPLICATION UNDER 37 CFR 1.53(b)

Assistant Commissioner for Patents
Box PATENT APPLICATION
Washington, D.C. 20231

Dear Sir:

This is a request for filing a **DIVISIONAL** application under 37 CFR 1.53(b) of pending prior application Serial No. 09/078,402, filed on May 13, 1998, entitled HUMAN APOPTOSIS ASSOCIATED PROTEINS.

1. __X__ Enclosed is a copy of the prior application, U.S. application Serial No. 09/078,402, filed May 13, 1998, including the oath or declaration as originally signed.

2. __X__ With regard to the requirement of 37 CFR 1.821(e) which requires that a copy of the Sequence Listing in computer readable form (CRF) be submitted, Applicants state that the computer readable form of the "Sequence Listing" for the instant divisional application is identical with that filed for Serial No. 09/078,402, filed May 13, 1998, to which priority is claimed. In accordance with 37 C.F.R. §1.821(e), please use the computer readable form filed with U.S. application Serial No. 09/078,402 as the computer readable form for the instant divisional application. It is understood that the Patent and Trademark Office will make the necessary change in application number and filing date for the computer readable form that will be used for the instant divisional application.

3. ___ Cancel in this application original claims _3, 6, 7, 9-12, 19, and 20_ of the prior application before calculating the filing fee. (At least one original independent claim must be retained for filing purposes.)

4. __X__ The inventors of the invention being claimed in this application are: Jennifer L. Hillman, Neil C. Corley, Karl J. Guegler, Chandra Patterson, and Mariah R. Baughn.

1

5. _X_ In accordance with 37 CFR 1.63(d), a copy of the originally signed declaration showing applicants' signatures as filed on July 16, 1998, is enclosed.

6. _X_ Amend the specification by inserting before the first line the sentence: "This application is a divisional application of U.S. application Serial Number 09/078,402, filed May 13, 1998."

7. _X_ The filing fee is calculated below:

Claims	Number Filed	Minus	Number Extra	Other Than Small Entity Rate	Fee	Basic Fee $760.00
Total Claims	20	-20	0	x $18		$ 0
Indep. Claims	2	-3	0	x $78		$ 0
__ Multiple Dependent Claim(s), if any				+ $260		$ 0

TOTAL FILING FEE $ 760.00

8. ___ An extension of time in the above-named prior application has been requested and the fees therefore have been authorized in said application.

9. _X_ Please charge Incyte Pharmaceuticals, Inc. Deposit Account No._09-0108_ in the amount of $_760.00_ .

The Commissioner is hereby authorized to charge any additional fees which may be required or credit any overpayment to Incyte Pharmaceuticals, Inc. Account No._09-0108_.

A **duplicate** copy of this Request is enclosed.

10. ___ New formal drawings are enclosed.

11. _X_ The prior application is assigned of record to _Incyte Pharmaceuticals, Inc._, recorded on July 21, 1998, at reel 9336/frame 0300.

12. ___ A preliminary amendment is enclosed.

13. _X_ Also enclosed: Return Postcard, Information Disclosure Statement, List of Cited References (1449), and copy of previously submitted Revocation of Power of Attorney and Appointment of New Attorneys.

14. __X__   The power of attorney of the prior application is to:

Narinder S. Banait	Reg. No. 43,482
Adam Warwick Bell	Reg. No. 43,490
Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Colette C. Muenzen	Reg. No. 39,784
Lynn E. Murry	Reg. No. 42,918
Danielle M. Pasqualone	Reg. No. 43,847
Susan K. Sather	Reg. No. 44,316
David G. Streeter	Reg. No. 43,168

___   The associate power of attorney in the prior application is to:

a. ___   An associate power of attorney is attached.

b. ___   Since the power does not appear in the original papers, a copy of the power in the prior application is enclosed.

c. __X__   Address all future correspondence to:

<div align="center">

INCYTE PHARMACEUTICALS, INC.
PATENT DEPARTMENT
3174 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555, Fax: (650) 849-8886

</div>

Date: _23 December 1999_   By: _____
Diana Hamlet-Cox
Reg. No. 33,302
Direct Dial Telephone: (650) 845-4639

___ Inventor(s)
___ Assignee of complete interest
__X__ Attorney or agent of record
___ Filed under 37 CFR 1.34(a)
       Registration number if acting under 37 CFR 1.34(a)_____ .

HUMAN APOPTOSIS ASSOCIATED PROTEINS

FIELD OF THE INVENTION

This invention relates to nucleic acid and amino acid sequences of human apoptosis associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of diseases associated with increased or decreased apoptosis.

5

BACKGROUND OF THE INVENTION

Normal development, growth, and homeostasis in multicellular organisms require a careful balance between the production and destruction of cells in tissues throughout the body. Cell division is a carefully coordinated process with numerous checkpoints and control mechanisms. These mechanisms are designed to regulate DNA replication and to prevent inappropriate or excessive cell proliferation. In contrast, apoptosis is the genetically controlled process by which unneeded or defective cells undergo programmed cell death. Unlike necrotic or injured cells, apoptotic cells are rapidly phagocytosed by neighboring cells or macrophages without leaking their potentially damaging contents into the surrounding tissue or triggering an inflammatory response.

Apoptotic cells undergo distinct morphological changes. Hallmarks of apoptosis include cell shrinkage, nuclear and cytoplasmic condensation, and alterations in plasma membrane topology. Biochemically, apoptotic cells are characterized by increased intracellular calcium concentration, fragmentation of chromosomal DNA into nucleosomal-length units, and expression of novel cell surface components.

Apoptotic events are part of the normal developmental programs of many multicellular organisms. Selective elimination of cells is as important for morphogenesis and tissue remodeling as cell proliferation and differentiation. Apoptosis is also a critical component of the immune response. Immune cells such as cytotoxic T-cells and natural killer cells induce apoptosis in tumor cells or virus-infected cells _in vitro_. Such activity _in vivo_ may halt the spread of virus or tumor proliferation. In addition, immune cells that fail to distinguish self molecules from foreign molecules must be eliminated to avoid an autoimmune response.

The molecular mechanisms of apoptosis are highly conserved, and many of the key protein regulators and effectors of apoptosis have been identified. Apoptosis generally proceeds in response to a signal which is transduced intracellularly, resulting in altered patterns of gene expression and protein activity. Signaling molecules such as hormones and cytokines are known to regulate apoptosis both positively and negatively. Transcription factors also play an important role in the onset of apoptosis. A number of downstream effector molecules, especially proteases, have been implicated in the degradation of cellular components and the proteolytic activation of other apoptotic effectors.

The rat ventral prostate (RVP) is a model system for the study of hormone-regulated apoptosis. RVP epithelial cells undergo apoptosis in response to androgen deprivation. Messenger RNA (mRNA) transcripts that are up-regulated in the apoptotic RVP have been identified. (Briehl, M. M. and Miesfeld, R. L. (1991) Mol. Endocrinol. 5:1381-1388.) One such transcript encodes RVP.1, the precise role of which in apoptosis has not been determined. The human homolog of RVP.1, hRVP1, is 89% identical to the rat protein. (Katahira, J. et al. (1997) J. Biol. Chem. 272:26652-26658.) hRVP1 is 220 amino acids in length and contains four transmembrane domains. hRVP1 is highly expressed in the lung, intestine, and liver. Interestingly, hRVP1 functions as a low affinity receptor for the Clostridium perfringens enterotoxin, a causative agent of diarrhea in humans and other animals.

Cytokine-mediated apoptosis plays an important role in hematopoiesis and the immune response. Myeloid cells, which are the stem cell progenitors of macrophages, neutrophils, erythrocytes, and other blood cells, proliferate in response to specific cytokines such as granulocyte/macrophage-colony stimulating factor (GM-CSF) and interleukin-3 (IL-3). When deprived of GM-CSF or IL-3, myeloid cells undergo apoptosis. The murine requiem (req) gene encodes a putative transcription factor required for this apoptotic response in the myeloid cell line FDCP-1. (Gabig, T. G. et al. (1994) J. Biol. Chem. 269:29515-29519.) The Req protein is 371 amino acids in length and contains a nuclear localization signal, a single Kruppel-type zinc finger, an acidic domain, and a cluster of four unique zinc-finger motifs enriched in cysteine and histidine residues involved in metal binding. Expression of req is not myeloid- or apoptosis-specific, suggesting that additional factors

regulate Req activity in myeloid cell apoptosis.

Tumor necrosis factor (TNF) and related cytokines induce apoptosis in lymphoid cells. (Reviewed in Nagata, S. (1997) Cell 88:355-365.) Binding of TNF to its receptor triggers a signal transduction pathway that results in the activation of a cascade of related

5   proteases, called caspases. One such caspase, ICE (Interleukin-1β converting enzyme), is a cysteine protease comprised of two large and two small subunits generated by ICE auto-cleavage. (Dinarello, C. A. (1994) FASEB J. 8:1314-1325.) ICE is expressed primarily in monocytes. ICE processes the cytokine precursor, interleukin-1β, into its active form, which plays a central role in acute and chronic inflammation, bone resorption, myelogenous

10   leukemia, and other pathological processes. ICE and related caspases cause apoptosis when overexpressed in transfected cell lines.

A final step in the apoptotic effector pathway is the fragmentation of nuclear DNA. Recently, a novel factor linking caspase activity to DNA fragmentation has been identified. (Xuesong, L. et al. (1997) Cell 89:175-184.) This factor, DNA fragmentation factor 45 (DFF-

15   45), is proteolytically activated by caspase and is required for DNA fragmentation. DFF-45 is 331 amino acids in length and exists in the cell as a heterodimer with a second uncharacterized factor. The amino acid sequence of DFF-45 indicates that it is not a nuclease, suggesting that DFF-45 may activate a downstream nuclease. In addition, mRNA encoding a protein related to DFF-45 has been isolated from mouse adipogenic cells.

20   (Danesch, U. et al. (1992) J. Biol. Chem. 267:7185-7193.) Expression of this mRNA is induced in steroid-treated, differentiating adipocytes. The predicted protein, FSP-27 (fat cell-specific, 27 kilodaltons), is highly basic with a predicted isoelectric point of 10.

Dysregulation of apoptosis has recently been recognized as a significant factor in the pathogenesis of many human diseases. For example, excessive cell survival caused by

25   decreased apoptosis can contribute to disorders related to cell proliferation and the immune response. Such disorders include cancer, autoimmune diseases, viral infections, and inflammation. In contrast, excessive cell death caused by increased apoptosis can lead to degenerative and immunodeficiency disorders such as AIDS, neurodegenerative diseases, and myelodysplastic syndromes. (Thompson, C.B. (1995) Science 267:1456-1462.)

30   The discovery of new human apoptosis associated proteins and the polynucleotides

encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, treatment, and prevention of diseases associated with increased or decreased apoptosis.

SUMMARY OF THE INVENTION

5

The invention is based on the discovery of new human apoptosis associated proteins (HAPOP), the polynucleotides encoding HAPOP, and the use of these compositions for the diagnosis, treatment, or prevention of diseases associated with increased or decreased apoptosis. The invention features substantially purified polypeptides, human apoptosis

10 associated proteins, referred to collectively as "HAPOP" and individually as "HAPOP-1," "HAPOP-2," "HAPOP-3," and "HAPOP-4." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ

15 ID NO:7.

The invention further provides a substantially purified variant having at least 90% amino acid identity to the amino acid sequences of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:7, or to a fragment of either of these sequences. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an

20 amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence

25 selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide

30 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ

ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7, as well as an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group

5   consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4,

10   SEQ ID NO:6, SEQ ID NO:8, a fragment of SEQ ID NO:2, a fragment of SEQ ID NO:4, a fragment of SEQ ID NO:6, and a fragment of SEQ ID NO:8. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, a

15   fragment of SEQ ID NO:2, a fragment of SEQ ID NO:4, a fragment of SEQ ID NO:6, and a fragment of SEQ ID NO:8, as well as an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, a fragment of SEQ ID NO:2, a fragment of SEQ ID NO:4, a fragment of SEQ

20   ID NO:6, and a fragment of SEQ ID NO:8.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a

25   fragment of SEQ ID NO:7. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a

30   fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7, the method comprising the steps

of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide encoding the polypeptide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7 in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7, as well as a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with increased or decreased apoptosis, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7.

The invention also provides a method for treating or preventing a disorder associated with increased or decreased apoptosis, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of the polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7.

The invention also provides a method for detecting a polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7 in a

6

biological sample containing nucleic acids, the method comprising the steps of: (a)
hybridizing the complement of the polynucleotide sequence encoding the polypeptide
comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1,
SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of

5   SEQ ID NO:3, a fragment of SEQ ID NO:5 and a fragment of SEQ ID NO:7 to at least one of
the nucleic acids of the biological sample, thereby forming a hybridization complex; and (b)
detecting the hybridization complex, wherein the presence of the hybridization complex
correlates with the presence of a polynucleotide encoding the polypeptide in the biological
sample. In one aspect, the nucleic acids of the biological sample are amplified by the

10   polymerase chain reaction prior to the hybridizing step.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A and 1B show the amino acid sequence alignments among HAPOP-2
(642272; SEQ ID NO:3), mouse FSP-27 (SWISS-PROT P56198; SEQ ID NO:9), and human

15   DFF-45 (GI 2065561; SEQ ID NO:10).

Figures 2A and 2B show the amino acid sequence alignment between HAPOP-3
(1453807; SEQ ID NO:5) and mouse Req (GI 606661; SEQ ID NO:11).

Figure 3 shows the amino acid sequence alignment between HAPOP-4 (2059022;
SEQ ID NO:7) and human hRVP1 (GI 2570129; SEQ ID NO:12).

20   These alignments were produced using the multisequence alignment program of
LASERGENE™ software (DNASTAR Inc, Madison WI).

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is

25   understood that this invention is not limited to the particular methodology, protocols, cell
lines, vectors, and reagents described, as these may vary. It is also to be understood that the
terminology used herein is for the purpose of describing particular embodiments only, and is
not intended to limit the scope of the present invention which will be limited only by the
appended claims.

30   It must be noted that as used herein and in the appended claims, the singular forms

"a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

5   Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are cited for the purpose
10   of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

15   **DEFINITIONS**

        "HAPOP," as used herein, refers to the amino acid sequences of substantially purified HAPOP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

20   The term "agonist," as used herein, refers to a molecule which, when bound to HAPOP, increases or prolongs the duration of the effect of HAPOP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HAPOP.

        An "allelic variant," as this term is used herein, is an alternative form of the gene
25   encoding HAPOP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of
30   these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HAPOP, as described herein, include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HAPOP or a polypeptide with at least one functional characteristic of HAPOP. Included within this definition are polymorphisms which may or

5   may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HAPOP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HAPOP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally

10   equivalent HAPOP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HAPOP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino

15   acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" or "amino acid sequence," as used herein, refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to

20   naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HAPOP which are preferably about 5 to about 15 amino acids in length, most preferably 14 amino acids, and which retain some biological activity or immunological activity of HAPOP. Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein

25   molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification," as used herein, relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction

30   (PCR) technologies well known in the art. (See, e.g., Dieffenbach, C.W. and G.S. Dveksler

(1995) <u>PCR Primer, a Laboratory Manual</u>, Cold Spring Harbor Press, Plainview, NY, pp.1-5.)

The term "antagonist," as it is used herein, refers to a molecule which, when bound to HAPOP, decreases the amount or the duration of the effect of the biological or immunological activity of HAPOP. Antagonists may include proteins, nucleic acids,

5   carbohydrates, antibodies, or any other molecules which decrease the effect of HAPOP.

As used herein, the term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, F(ab')$_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind HAPOP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing

10   antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

15   The term "antigenic determinant," as used herein, refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may

20   compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense," as used herein, refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or

25   transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

As used herein, the term "biologically active," refers to a protein having structural,

30   regulatory, or biochemical functions of a naturally occurring molecule. Likewise,

"immunologically active" refers to the capability of the natural, recombinant, or synthetic HAPOP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" or "complementarity," as used herein, refer to the natural

5 binding of polynucleotides under permissive salt and temperature conditions by base pairing. For example, the sequence "A-G-T" binds to the complementary sequence "T-C-A." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between

10 nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" or a "composition

15 comprising a given amino acid sequence," as these terms are used herein, refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation, an aqueous solution, or a sterile composition. Compositions comprising polynucleotide sequences encoding HAPOP or fragments of HAPOP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be

20 associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts, e.g., NaCl, detergents, e.g.,sodium dodecyl sulfate (SDS), and other components, e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.

"Consensus sequence," as used herein, refers to a nucleic acid sequence which has

25 been resequenced to resolve uncalled bases, extended using XL-PCR™ (Perkin Elmer, Norwalk, CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW™ Fragment Assembly system (GCG, Madison, WI). Some sequences have been both extended and assembled to produce the consensus

30 sequence.

11

As used herein, the term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HAPOP, by Northern analysis is indicative of the presence of nucleic acids encoding HAPOP in a sample, and thereby correlates with expression of the transcript

5   from the polynucleotide encoding HAPOP.

A "deletion," as the term is used herein, refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative," as used herein, refers to the chemical modification of a

10   polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one

15   biological or immunological function of the polypeptide from which it was derived.

The term "similarity," as used herein, refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially

20   similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced

25   stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the

30   absence of non-specific binding, the substantially similar sequence or probe will not hybridize

to the second non-complementary target sequence.

The phrases "percent identity" or "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MegAlign™ program

5   (DNASTAR, Inc., Madison WI). The MegAlign™ program can create alignments between two or more sequences according to different methods, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences,

10   e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also

15   be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs), as described herein, are linear

20   microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance. (See, e.g., Harrington, J.J. et al. (1997) Nat Genet. 15:345-355.)

The term "humanized antibody," as used herein, refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the

25   antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization," as the term is used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

As used herein, the term "hybridization complex" refers to a complex formed between

30   two nucleic acid sequences by virtue of the formation of hydrogen bonds between

complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" or "addition," as used herein, refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray," as used herein, refers to an arrangement of distinct polynucleotides arrayed on a substrate, e.g., paper, nylon or any other type of membrane, filter, chip, glass slide, or any other suitable solid support.

The terms "element" or "array element" as used herein in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate," as it appears herein, refers to a change in the activity of HAPOP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HAPOP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" or "operably linked," as used herein, refer to

14

functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are

5   not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

        The term "oligonucleotide," as used herein, refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a

10   hybridization assay or microarray. As used herein, the term "oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

        "Peptide nucleic acid" (PNA), as used herein, refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length

15   linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell. (See, e.g., Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63.)

        The term "sample," as used herein, is used in its broadest sense. A biological sample

20   suspected of containing nucleic acids encoding HAPOP, or fragments thereof, or HAPOP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a solid support; a tissue; a tissue print; etc.

        As used herein, the terms "specific binding" or "specifically binding" refer to that

25   interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody

30   will reduce the amount of labeled A that binds to the antibody.

As used herein, the term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent (e.g., formamide), temperature, and other conditions well known in the art. In particular, stringency can be
5   increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium
10   citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying
15   additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment,
20   hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.
25   The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about
30   15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash

steps will ordinarily include temperature of at least about 25°C, more preferably of at least

about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash

steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more

preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium

5   citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15

mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these

conditions will be readily apparent to those skilled in the art.

The term "substantially purified," as used herein, refers to nucleic acid or amino acid

sequences that are removed from their natural environment and are isolated or separated, and

10   are at least about 60% free, preferably about 75% free, and most preferably about 90% free

from other components with which they are naturally associated.

A "substitution," as used herein, refers to the replacement of one or more amino acids

or nucleotides by different amino acids or nucleotides, respectively.

"Transformation," as defined herein, describes a process by which exogenous DNA

15   enters and changes a recipient cell. Transformation may occur under natural or artificial

conditions according to various methods well known in the art, and may rely on any known

method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic

host cell. The method for transformation is selected based on the type of host cell being

transformed and may include, but is not limited to, viral infection, electroporation, heat

20   shock, lipofection, and particle bombardment. The term "transformed" cells includes stably

transformed cells in which the inserted DNA is capable of replication either as an

autonomously replicating plasmid or as part of the host chromosome, as well as transiently

transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of HAPOP, as used herein, refers to an amino acid sequence that is altered

25   by one or more amino acids. The variant may have "conservative" changes, wherein a

substituted amino acid has similar structural or chemical properties (e.g., replacement of

leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g.,

replacement of glycine with tryptophan). Analogous minor variations may also include

amino acid deletions or insertions, or both. Guidance in determining which amino acid

30   residues may be substituted, inserted, or deleted without abolishing biological or

immunological activity may be found using computer programs well known in the art, for example, LASERGENE™ software.

THE INVENTION

5   The invention is based on the discovery of new human apoptosis associated proteins (HAPOP), the polynucleotides encoding HAPOP, and the use of these compositions for the diagnosis, treatment, or prevention of diseases associated with increased or decreased apoptosis.

      Nucleic acids encoding the HAPOP-1 of the present invention were first identified in

10   Incyte Clone 157658 from the promonocyte cDNA library (THP1PLB02) using a computer search, e.g., BLAST, for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 157658 (THP1PLB02), 2451287 (ENDANOT01), and 1241829 (LUNGNOT03).

15   In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1. HAPOP-1 is 480 amino acids in length and has one potential N-glycosylation site at N262; eight potential casein kinase II phosphorylation sites at S51, S130, S193, S227, S238, S371, S411, and S458; and five potential protein kinase C phosphorylation sites at S181, S193, S428, S458, and T470. Both BLOCKS and PRINTS

20   analyses indicate that regions of HAPOP-1 show similarity to conserved motifs in ICE. These regions extend from P250 to R303, from F308 to G316, and from S344 to S363. A region of unique sequence in HAPOP-1 from about amino acid 203 to about amino acid 212 is encoded by a fragment of SEQ ID NO:2 from about nucleotide 1218 to about nucleotide 1247. Northern analysis shows the expression of this sequence in various libraries, at least

25   61% of which are associated with proliferating or cancerous tissue and at least 35% of which are associated with the immune response. In particular, 26% of the libraries expressing HAPOP-1 are derived from reproductive tissue.

      Nucleic acids encoding the HAPOP-2 of the present invention were first identified in Incyte Clone 642272 from the breast cDNA library (BRSTNOT03) using a computer search,

30   e.g., BLAST, for amino acid sequence alignments. A consensus sequence, SEQ ID NO:4, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte

Clones 642272 (BRSTNOT03), 1954870 (CONNNOT01), 2904084 (DRGCNOT01), and 1965381 (BRSTNOT04).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:3. HAPOP-2 is 238 amino acids in length and has four

5 potential casein kinase II phosphorylation sites at S48, S62, T90, and T164 and four potential protein kinase C phosphorylation sites at S21, S36, S53, and S131. As shown in Figures 1A and 1B, HAPOP-2 has chemical and structural similarity with mouse FSP-27 (SWISS-PROT P56198; SEQ ID NO:9) and human DFF-45 (GI 2065561; SEQ ID NO:10). In particular, HAPOP-2 and FSP-27 share 79% identity, and HAPOP-2 and DFF-45 share 18% identity. In

10 addition, the potential phosphorylation sites at S48, S62, T164, and S131 in HAPOP-2 are conserved in FSP-27, and the potential phosphorylation site at T90 in HAPOP-2 is conserved in both FSP-27 and DFF-45. Like FSP-27, HAPOP-2 is a basic protein with a predicted isoelectric point of 8.8. A region of unique sequence in HAPOP-2 from about amino acid 31 to about amino acid 40 is encoded by a fragment of SEQ ID NO:4 from about nucleotide 229

15 to about nucleotide 258. Northern analysis shows the expression of this sequence in various libraries, at least 55% of which are associated with proliferating or cancerous tissue and at least 48% of which are associated with the immune response. In particular, 35% of the libraries expressing HAPOP-2 are derived from reproductive tissue, and 30% are derived from gastrointestinal tissue.

20 Nucleic acids encoding the HAPOP-3 of the present invention were first identified in Incyte Clone 1453807 from the penis tumor cDNA library (PENITUT01) using a computer search, e.g., BLAST, for amino acid sequence alignments. A consensus sequence, SEQ ID NO:6, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 1453807 (PENITUT01), 2690812 (LUNGNOT23), 1392458 (THYRNOT03),

25 2344196 (TESTTUT02), 1995291 and 899939 (BRSTTUT03), 1495966 (PROSNON01), 2083186 (UTRSNOT08), and 3239041 (COLAUCT01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:5. HAPOP-3 is 410 amino acids in length and has one potential N-glycosylation site at N237; one potential cAMP- and cGMP-dependent protein kinase

30 phosphorylation site at S182; eight potential casein kinase II phosphorylation sites at S28,

19

S182, S209, S239, S243, S247, S314, and S323; eight potential protein kinase C phosphorylation sites at S14, S28, T100, T157, T264, S301, S382, and T396; and two potential tyrosine kinase phosphorylation sites at Y137 and Y188. As shown in Figures 2A and 2B, HAPOP-3 has chemical and structural similarity with mouse Req (GI 606661; SEQ ID NO:11). HAPOP-3 and Req share 20% identity. In particular, the region of HAPOP-3 from I291 to Q391 is 44% identical to the corresponding region of Req, which comprises the cluster of four unique zinc finger motifs. All 16 cysteine and histidine residues required for metal binding in Req are conserved in HAPOP-3. A region of unique sequence in HAPOP-3 from about amino acid 36 to about amino acid 45 is encoded by a fragment of SEQ ID NO:6 from about nucleotide 181 to about nucleotide 210. Northern analysis shows the expression of this sequence in various libraries, at least 67% are associated with proliferating or cancerous tissue and at least 33% are associated with the immune response. In particular, 35% of the libraries expressing HAPOP-3 are derived from reproductive tissue, and 27% are derived from gastrointestinal tissue.

Nucleic acids encoding the HAPOP-4 of the present invention were first identified in Incyte Clone 2059022 from the ovarian cDNA library (OVARNOT03) using a computer search, e.g., BLAST, for amino acid sequence alignments. A consensus sequence, SEQ ID NO:8, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 2059022 (OVARNOT03), 1864717 (PROSNOT19), 814715 (OVARTUT01), 772645 (COLNCRT01), 1317994 (BLADNOT04), 1434058 (BEPINON01), and 1222796 (COLNTUT02).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:12. HAPOP-4 is 211 amino acids in length and has two potential N-glycosylation sites at N189 and N205; one potential casein kinase II phosphorylation site at S206; two potential protein kinase C phosphorylation sites at S63 and S206; and a potential signal peptide sequence from M1 to about C24. As shown in Figure 3, HAPOP-4 has chemical and structural similarity with hRVP1 (GI 2570129; SEQ ID NO:12). In particular, HAPOP-4 and hRVP1 share 44% identity. Hydrophobicity plots demonstrate that the four transmembrane domains of hRVP1 are well conserved in HAPOP-4 from about amino acid 1 to about amino acid 29; from about amino acid 78 to about amino acid 103;

20

from about amino acid 120 to about amino acid 140; and from amino acid 162 to about amino acid 186. A region of unique sequence in HAPOP-4 from about amino acid 30 to about amino acid 39 is encoded by a fragment of SEQ ID NO:8 from about nucleotide 520 to about nucleotide 549. Northern analysis shows the expression of this sequence in various libraries,

5   at least 70% are associated with proliferating or cancerous tissue and at least 30% are associated with the immune response. In particular, 38% of the libraries expressing HAPOP-4 are derived from reproductive tissue, and 24% are derived from gastrointestinal tissue.

The invention also encompasses HAPOP variants. A preferred HAPOP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least

10   about 95% amino acid sequence identity to the HAPOP amino acid sequence, and which contains at least one functional or structural characteristic of HAPOP.

The invention also encompasses polynucleotides which encode HAPOP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID NO:2, which encodes HAPOP-1. In a further embodiment, the invention

15   encompasses a polynucleotide sequence comprising the sequence of SEQ ID NO:4, which encodes HAPOP-2. In another embodiment, the invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID NO:6, which encodes HAPOP-3. In still another embodiment, the invention encompasses a polynucleotide sequence comprising the sequence of SEQ ID NO:8, which encodes HAPOP-4.

20   The invention also encompasses a variant of a polynucleotide sequence encoding HAPOP. In particular, such a variant polynucleotide sequence will have at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HAPOP. A particular aspect of the invention encompasses a variant of SEQ ID NO:2 which has at least about 80%, more

25   preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to SEQ ID NO:2. The invention further encompasses a polynucleotide variant of SEQ ID NO:4 having at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to SEQ ID NO:4. The invention further encompasses a polynucleotide variant of SEQ ID NO:6 having at least about

30   80%, more preferably at least about 90%, and most preferably at least about 95%

21

polynucleotide sequence identity to SEQ ID NO:6. The invention further encompasses a polynucleotide variant of SEQ ID NO:8 having at least about 80%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to SEQ ID NO:8. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HAPOP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HAPOP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring HAPOP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HAPOP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HAPOP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HAPOP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HAPOP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HAPOP and HAPOP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HAPOP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, a fragment of SEQ ID NO:2, a fragment of SEQ ID NO:4, a fragment of SEQ ID NO:6, or a fragment of SEQ ID NO:8,

5   under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.)

Methods for DNA sequencing are well known and generally available in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical

10   Corp., Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE™ Amplification System (GIBCO BRL, Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; MJ Research, Watertown,

15   MA) and the ABI Catalyst and 373 and 377 DNA Sequencers (Perkin Elmer).

The nucleic acid sequences encoding HAPOP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify

20   unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A

25   third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to

30   retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991)

23

Nucleic Acids Res. 19:3055-306). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk genomic DNA (Clontech, Palo Alto, CA). This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such

5   as OLIGO™ 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often

10   include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In

15   particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., Genotyper™ and Sequence Navigator™, Perkin Elmer), and the entire process from loading of samples to computer

20   analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HAPOP may be cloned in recombinant DNA molecules that direct

25   expression of HAPOP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HAPOP.

The nucleotide sequences of the present invention can be engineered using methods

30   generally known in the art in order to alter HAPOP-encoding sequences for a variety of

24

purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may

5   be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HAPOP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucl. Acids Res. Symp. Ser. 215-223, and Horn, T. et al. (1980) Nucl. Acids Res.

10   Symp. Ser. 225-232.) Alternatively, HAPOP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A Peptide Synthesizer (Perkin Elmer). Additionally, the amino acid sequence of HAPOP, or any part thereof, may be

15   altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid

20   analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman and Co., New York, NY.)

In order to express a biologically active HAPOP, the nucleotide sequences encoding HAPOP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of

25   the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HAPOP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HAPOP. Such signals include the ATG initiation

30   codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding

HAPOP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon

5   should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct

10   expression vectors containing sequences encoding HAPOP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview, NY, ch. 4, 8, and 16-17; and Ausubel, F.M. et al. (1995, and periodic supplements) Current

15   Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HAPOP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with

20   viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus (CaMV) or tobacco mosaic virus (TMV)) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected

25   depending upon the use intended for polynucleotide sequences encoding HAPOP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HAPOP can be achieved using a multifunctional E. coli vector such as Bluescript® (Stratagene) or pSport1™ plasmid (GIBCO BRL). Ligation of sequences encoding HAPOP into the vector's multiple cloning site disrupts the lacZ gene, allowing a colorimetric

30   screening procedure for identification of transformed bacteria containing recombinant

molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HAPOP are needed, e.g. for the production of

5   antibodies, vectors which direct high level expression of HAPOP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HAPOP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition,

10   such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, supra; and Grant et al. (1987) Methods Enzymol. 153:516-54; Scorer, C. A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HAPOP. Transcription of sequences encoding HAPOP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV

15   used alone or in combination with the omega leader sequence from TMV. (Takamatsu, N. (1987) EMBO J. 6:307-311.) Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991)

20   Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., Hobbs, S. or Murry, L.E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, NY; pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In

25   cases where an adenovirus is used as an expression vector, sequences encoding HAPOP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HAPOP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659.) In addition,

30   transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to

increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes.

For long term production of recombinant proteins in mammalian systems, stable expression of HAPOP in cell lines is preferred. For example, sequences encoding HAPOP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* or *apr⁻* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; and Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14; and Murry, supra.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP) (Clontech, Palo Alto, CA), ß glucuronidase and its substrate ß-D-glucuronoside, or luciferase and its substrate luciferin may be used. These markers can be

used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of

5   interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HAPOP is inserted within a marker gene sequence, transformed cells containing sequences encoding HAPOP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HAPOP under the control of a single promoter. Expression of the marker gene in

10   response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HAPOP and that express HAPOP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which

15   include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HAPOP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays

20   (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HAPOP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Section IV; Coligan, J. E. et al. (1997 and

25   periodic supplements) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY; and Maddox, D.E. et al. (1983) J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides

30   encoding HAPOP include oligolabeling, nick translation, end-labeling, or PCR amplification

using a labeled nucleotide. Alternatively, the sequences encoding HAPOP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled

5   nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Pharmacia & Upjohn (Kalamazoo, MI), Promega (Madison, WI), and U.S. Biochemical Corp. (Cleveland, OH). Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors,

10   magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HAPOP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in

15   the art, expression vectors containing polynucleotides which encode HAPOP may be designed to contain signal sequences which direct secretion of HAPOP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such

20   modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38),

25   are available from the American Type Culture Collection (ATCC, Bethesda, MD) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HAPOP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a

30   chimeric HAPOP protein containing a heterologous moiety that can be recognized by a

30

commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HAPOP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein

5   (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and

10   polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HAPOP encoding sequence and the heterologous protein sequence, so that HAPOP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel, F. M. et al. (1995 and periodic

15   supplements) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch 10. A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HAPOP may be achieved in vitro using the TNT™ rabbit reticulocyte lysate or wheat germ extract systems

20   (Promega, Madison, WI). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of HAPOP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-

25   60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of HAPOP may be synthesized separately and then combined to produce the full length molecule.

30   **THERAPEUTICS**

Chemical and structural similarity exists between HAPOP-1 and conserved motifs in

ICE. In addition, HAPOP-1 is expressed in proliferating, immune, and reproductive tissue.

Chemical and structural similarity exists among HAPOP-2, mouse FSP-27 (SWISS-PROT P56198), and human DFF-45 (GI 2065561). Chemical and structural similarity exists between HAPOP-3 and mouse Req (GI 606661) and between HAPOP-4 and human hRVP1

5 (GI 2570129). In addition, HAPOP-2, HAPOP-3, and HAPOP-4 are expressed in proliferating, immune, reproductive and gastrointestinal tissue.

Therefore, HAPOP appears to play a role in disorders associated with increased or decreased apoptosis, particularly those disorders associated with cell proliferation and the immune, reproductive and gastrointestinal systems.

10 Therefore, in one embodiment, HAPOP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with increased or decreased apoptosis. Such disorders can include, but are not limited to, cell proliferative disorders such as atheroscleosis, arteriosclerosis, and cancers such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal

15 gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; immune disorders such as infections, rheumatoid arthritis, dermatitis, inflammation, systemic lupus erythematosus and other autoimmune diseases; reproductive disorders such as endometriosis, endometrial and ovarian

20 tumors, breast cancer, fibrocystic breast disease, testicular cancer, prostate cancer, and gynecomastia; and gastrointestinal disorders such as esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease, cholecystitis, infections of the intestinal tract, pancreatitis, pancreatic carcinoma, cirrhosis, hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease.

25 In another embodiment, a vector capable of expressing HAPOP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with increased or decreased apoptosis, including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HAPOP in conjunction with a suitable pharmaceutical carrier may be administered to

30 a subject to treat or prevent a disorder associated with increased or decreased apoptosis,

including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HAPOP may be administered to a subject to treat or prevent a disorder associated with increased or decreased apoptosis, including, but not limited to, those listed above.

5   In a further embodiment, an antagonist of HAPOP may be administered to a subject to treat or prevent a disorder associated with increased or decreased apoptosis. Such disorders may include, but are not limited to, those discussed above. In one aspect, an antibody which specifically binds HAPOP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express

10   HAPOP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HAPOP may be administered to a subject to treat or prevent a disorder associated with increased or decreased apoptosis, including, but not limited to, those described above.

15   In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act

20   synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HAPOP may be produced using methods which are generally known in the art. In particular, purified HAPOP may be used to produce antibodies or to screen

25   libraries of pharmaceutical agents to identify those which specifically bind HAPOP. Antibodies to HAPOP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for

30   therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HAPOP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include,

5 but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies

10 to HAPOP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HAPOP amino acids may be fused with those of another protein, such as

15 KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HAPOP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature

20 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with

25 appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HAPOP-specific single chain antibodies. Antibodies with related

30 specificity, but of distinct idiotypic composition, may be generated by chain shuffling from

random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:10134-10137.)

Antibodies may also be produced by inducing _in vivo_ production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding

5   reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; and Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HAPOP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by

10   reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the

15   desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HAPOP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HAPOP

20   epitopes is preferred, but a competitive binding assay may also be employed. (Maddox, _supra_.)

In another embodiment of the invention, the polynucleotides encoding HAPOP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HAPOP may be used in situations in which it

25   would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HAPOP. Thus, complementary molecules or fragments may be used to modulate HAPOP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations

30   along the coding or control regions of sequences encoding HAPOP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences

5   complementary to the polynucleotides encoding HAPOP. (See, e.g., Sambrook, supra; and Ausubel, supra.)

Genes encoding HAPOP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HAPOP. Such constructs may be used to introduce untranslatable sense or

10   antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing

15   complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HAPOP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to

20   open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of

25   mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules

30   may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding

36

HAPOP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HAPOP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic

37

amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nature Biotechnology 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits,

5   monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HAPOP, antibodies to HAPOP, and mimetics, agonists, antagonists, or

10   inhibitors of HAPOP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

15   The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain

20   suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

25   Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

30   Pharmaceutical preparations for oral use can be obtained through combining active

compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants;

5   cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

10   Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

15   Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable

20   liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium

25   carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or

30   agents to increase the solubility of the compounds and allow for the preparation of highly

concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acid. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HAPOP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HAPOP or fragments thereof, antibodies of HAPOP, and agonists, antagonists or inhibitors of HAPOP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with

experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of therapeutic to toxic effects is the therapeutic index, and it can be expressed as the ED_{50}/LD_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are

5 preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

10 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s),

15 reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular

20 dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

25 **DIAGNOSTICS**

In another embodiment, antibodies which specifically bind HAPOP may be used for the diagnosis of disorders characterized by expression of HAPOP, or in assays to monitor patients being treated with HAPOP or agonists, antagonists, or inhibitors of HAPOP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described

30 above for therapeutics. Diagnostic assays for HAPOP include methods which utilize the antibody and a label to detect HAPOP in human body fluids or in extracts of cells or tissues.

41

The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HAPOP, including ELISAs, RIAs, and FACS,

5   are known in the art and provide a basis for diagnosing altered or abnormal levels of HAPOP expression. Normal or standard values for HAPOP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HAPOP under conditions suitable for complex formation The amount of standard complex formation may be quantitated by various methods, preferably by

10   photometric means. Quantities of HAPOP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HAPOP may be used for diagnostic purposes. The polynucleotides which may be used include

15   oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HAPOP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HAPOP, and to monitor regulation of HAPOP levels during therapeutic intervention.

20   In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HAPOP or closely related molecules may be used to identify nucleic acid sequences which encode HAPOP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of

25   the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HAPOP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HAPOP encoding sequences. The

30   hybridization probes of the subject invention may be DNA or RNA and may be derived from

42

the sequences of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, or from

genomic sequences including promoters, enhancers, and introns of the HAPOP gene.

Means for producing specific hybridization probes for DNAs encoding HAPOP

include the cloning of polynucleotide sequences encoding HAPOP or HAPOP derivatives

5   into vectors for the production of mRNA probes. Such vectors are known in the art, are

commercially available, and may be used to synthesize RNA probes in vitro by means of the

addition of the appropriate RNA polymerases and the appropriate labeled nucleotides.

Hybridization probes may be labeled by a variety of reporter groups, for example, by

radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled

10   to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HAPOP may be used for the diagnosis of a

disorder associated with expression of HAPOP. Examples of such a disorder include, but are

not limited to, a disorder associated with increased or decreased apoptosis, including cell

proliferative disorders such as atheroscleosis, arteriosclerosis, and cancers such as

15   adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and,

in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix,

gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas,

parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus;

immune disorders such as infections, rheumatoid arthritis, dermatitis, inflammation, systemic

20   lupus erythematosus and other autoimmune diseases; reproductive disorders such as

endometriosis, endometrial and ovarian tumors, breast cancer, fibrocystic breast disease,

testicular cancer, prostate cancer, and gynecomastia; and gastrointestinal disorders such as

esophagitis, esophageal carcinoma, gastritis, gastric carcinoma, inflammatory bowel disease,

cholecystitis, infections of the intestinal tract, pancreatitis, pancreatic carcinoma, cirrhosis,

25   hepatitis, hepatoma, colitis, colonic carcinoma, and Crohn's disease. The polynucleotide

sequences encoding HAPOP may be used in Southern or Northern analysis, dot blot, or other

membrane-based technologies; in PCR technologies; in dipstick, pin, and ELISA assays; and

in microarrays utilizing fluids or tissues from patients to detect altered HAPOP expression.

Such qualitative or quantitative methods are well known in the art.

30   In a particular aspect, the nucleotide sequences encoding HAPOP may be useful in

43

assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HAPOP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the

5   signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding HAPOP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the

10   treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HAPOP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HAPOP, under conditions

15   suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the

20   presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period

25   ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to

30   employ preventative measures or aggressive treatment earlier thereby preventing the

development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HAPOP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment

5 of a polynucleotide encoding HAPOP, or a fragment of a polynucleotide complementary to the polynucleotide encoding HAPOP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HAPOP include

10 radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; and Duplaa, C. et al. (1993) Anal. Biochem. 229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or

15 colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This

20 information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996)

25 Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HAPOP

30 may be used to generate hybridization probes useful in mapping the naturally occurring

45

genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries.

5 (See, e.g., Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, R.A. (ed.) Molecular Biology and Biotechnology, VCH Publishers New

10 York, NY, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HAPOP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to

15 detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a

20 particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area

25 may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HAPOP, its catalytic or immunogenic

30 fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of

a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HAPOP and the agent being tested may be measured.

5   Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with HAPOP, or fragments thereof, and washed. Bound

10   HAPOP is then detected by methods well known in the art. Purified HAPOP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which

15   neutralizing antibodies capable of binding HAPOP specifically compete with a test compound for binding HAPOP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HAPOP.

In additional embodiments, the nucleotide sequences which encode HAPOP may be used in any molecular biology techniques that have yet to be developed, provided the new

20   techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

25   **EXAMPLES**

I.   cDNA Library Construction

The BRSTNOT03 cDNA library was constructed from breast tissue removed from a 54-year-old Caucasian female during a bilateral radical mastectomy. Pathology for the associated tumor tissue indicated residual invasive grade 3 mammary ductal adenocarcinoma.

30   The remaining breast parenchyma exhibited proliferative fibrocystic changes without atypia.

One of 10 axillary lymph nodes had metastatic tumor, as a microscopic intranodal focus.

The PENITUT01 cDNA library was constructed from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin
5 and extending onto the glans penis. The tumor involved the glans but did not involve Buck's fascia or corpora cavernosa. Patient history included benign neoplasm of the large bowel and atherosclerotic coronary artery disease. Family history included malignant neoplasm, chronic lymphocytic leukemia, and chronic liver disease.

The OVARNOT03 cDNA library was constructed from ovarian tissue removed from
10 a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology for the associated tumor tissue indicated grade 2 mucinous cystadenocarcinoma. Patient history included viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, cerebrovascular disease, breast cancer, and uterine cancer.

Frozen tissue from each of the above sources was homogenized and lysed in
15 guanidinium isothiocyanate solution using a Brinkmann Homogenizer Polytron PT-3000 (Brinkmann Instruments, Westbury, NY). The lysate was centrifuged over a CsCl cushion to isolate RNA. The RNA was extracted with phenol, precipitated with sodium acetate and ethanol, resuspended in RNase-free water, and treated with DNase. The RNA was re-extracted as necessary with acid phenol and reprecipitated.

20 From each RNA preparation, poly (A+) RNA was isolated using the Qiagen Oligotex kit (QIAGEN, Chatsworth, CA). Poly (A+) RNA was used for cDNA synthesis and construction of each cDNA library according to the recommended protocols in the SuperScript plasmid system (Catalog #18248-013; Gibco BRL). The cDNAs were size-fractionated on a Sepharose CL4B column (Catalog #275105-01, Pharmacia, Piscataway, NJ).
25 Those cDNAs exceeding 400 bp were ligated into pINCY 1 (Incyte) (PENITUT01) or pSPORT 1 (Gibco BRL) (BRSTNOT03, OVARNOT03). Recombinant plasmids were transformed into DH5α™ competent cells (Catalog #18258-012, Gibco BRL).

The THP1PLB02 cDNA library was constructed by reamplification of the THP1PLB01 library, which was custom constructed by Stratagene (La Jolla, CA). THP-1 is a
30 human promonocyte line derived from the peripheral blood of a 1-year-old male with acute

48

monocytic leukemia. THP-1 cells were activated by culturing in the presence of phorbol ester (PMA) and lipopolysaccharide (LPS). Poly (A+) RNA was isolated from activated THP-1 cells, and cDNA synthesis was initiated using oligo(dT) and random primers. Double-stranded cDNA was cloned using the Lambda UniZAP vector system (Stratagene).

5

II. Isolation and Sequencing of cDNA Clones
PENITUT01, BRSTNOT03, AND OVARNOT03

Plasmid DNA was released from the cells and purified using the REAL Prep 96 plasmid kit (Catalog #26173, QIAGEN Inc) (PENITUT01 and BRSTNOT03) or the Miniprep

10 kit (Catalog #77468, Advanced Genetic Technologies Corporation, Gaithersburg, MD) (OVARNOT03). The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Catalog #22711, Gibco BRL) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after the cultures were incubated for 19-24 hours, the cells were lysed with 0.3-0.6 ml of lysis buffer; and 3)

15 following precipitation, plasmids were stored at 4°C.

THP1PLB02

Recombinant pBluescript® phagemid vectors containing cloned cDNAs were recovered by in vivo excision as described by Stratagene. Phagemid DNA was purified using the QIAwell-8 plasmid purification system (QIAGEN, Chatsworth, CA) or the Magic

20 Minipreps™ DNA purification system (Catalog #A7100, Promega, Madison, WI).

cDNAs from all four libraries were sequenced by the method of Sanger et al. (1975, J. Mol. Biol. 94:441f). Chain termination products were electrophoresed on urea-polyacrylamide gels and detected by autoradiography. Alternatively, high-throughput

25 methods for sample preparation utilized the Catalyst 800 or the Hamilton Micro Lab 2200 (Hamilton, Reno, NV) in combination with Peltier Thermal Cyclers (PTC200 from MJ Research, Watertown, MA). cDNAs were sequenced using the Applied Biosystems 373 or 377 DNA sequencing systems in conjunction with fluorescence detection methods; and the reading frames were determined.

30

III. Similarity Searching of cDNA Clones and Their Deduced Proteins

The nucleotide sequences and/or amino acid sequences of the Sequence Listing were used to query sequences in the GenBank, SwissProt, BLOCKS, and Pima II databases. These databases, which contain previously identified and annotated sequences, were searched for

5 regions of similarity using BLAST (Basic Local Alignment Search Tool). (See, e.g., Altschul, S.F. (1993) J. Mol. Evol 36:290-300; and Altschul et al. (1990) J. Mol. Biol. 215:403-410.)

BLAST produced alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST was

10 especially useful in determining exact matches or in identifying homologs which may be of prokaryotic (bacterial) or eukaryotic (animal, fungal, or plant) origin. Other algorithms could have been used when dealing with primary sequence patterns and secondary structure gap penalties. (See, e.g., Smith, T. et al. (1992) Protein Engineering 5:35-51.) The sequences disclosed in this application have lengths of at least 49 nucleotides and have no more than

15 12% uncalled bases (where N is recorded rather than A, C, G, or T).

The BLAST approach searched for matches between a query sequence and a database sequence. BLAST evaluated the statistical significance of any matches found, and reported only those matches that satisfy the user-selected threshold of significance. In this application, threshold was set at 10^{-25} for nucleotides and 10^{-8} for peptides.

20 Incyte nucleotide sequences were searched against the GenBank databases for primate (pri), rodent (rod), and other mammalian sequences (mam), and deduced amino acid sequences from the same clones were then searched against GenBank functional protein databases, mammalian (mamp), vertebrate (vrtp), and eukaryote (eukp), for similarity.

Additionally, sequences identified from cDNA libraries may be analyzed to identify

25 those gene sequences encoding conserved protein motifs using an appropriate analysis program, e.g., BLOCKS. BLOCKS is a weighted matrix analysis algorithm based on short amino acid segments, or blocks, compiled from the PROSITE database. (Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.) The BLOCKS algorithm is useful for classifying genes with unknown functions. (Henikoff S. And Henikoff G.J., Nucleic Acids Research

30 (1991) 19:6565-6572.) Blocks, which are 3-60 amino acids in length, correspond to the most

highly conserved regions of proteins. The BLOCKS algorithm compares a query sequence with a weighted scoring matrix of blocks in the BLOCKS database. Blocks in the BLOCKS database are calibrated against protein sequences with known functions from the SWISS-PROT database to determine the stochastic distribution of matches. Similar databases such as

5 PRINTS, a protein fingerprint database, are also searchable using the BLOCKS algorithm. (Attwood, T. K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.) PRINTS is based on non-redundant sequences obtained from sources such as SWISS-PROT, GenBank, PIR, and NRL-3D.

The BLOCKS algorithm searches for matches between a query sequence and the

10 BLOCKS or PRINTS database and evaluates the statistical significance of any matches found. Matches from a BLOCKS or PRINTS search can be evaluated on two levels, local similarity and global similarity. The degree of local similarity is measured by scores, and the extent of global similarity is measured by score ranking and probability values. A score of 1000 or greater for a BLOCKS match of highest ranking indicates that the match falls within

15 the 0.5 percentile level of false positives when the matched block is calibrated against SWISS-PROT. Likewise, a probability value of less than 1.0×10^{-3} indicates that the match would occur by chance no more than one time in every 1000 searches. Only those matches with a cutoff score of 1000 or greater and a cutoff probability value of 1.0×10^{3} or less are considered in the functional analyses of the protein sequences in the Sequence Listing.

20 Nucleic and amino acid sequences of the Sequence Listing may also be analyzed using PFAM. PFAM is a Hidden Markov Model (HMM) based protocol useful in protein family searching. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Cur. Opin. Str. Biol. 6:361-365.)

The PFAM database contains protein sequences of 527 protein families gathered from

25 publicly available sources, e.g., SWISS-PROT and PROSITE. PFAM searches for well characterized protein domain families using two high-quality alignment routines, seed alignment and full alignment. (See, e.g., Sonnhammer, E.L.L. et al. (1997) Proteins 28:405-420.) The seed alignment utilizes the hmmls program, a program that searches for local matches, and a non-redundant set of the PFAM database. The full alignment utilizes the

30 hmmfs program, a program that searches for multiple fragments in long sequences, e.g., repeats and motifs, and all sequences in the PFAM database. A result or score of 100 "bits"

can signify that it is 2^{100}-fold more likely that the sequence is a true match to the model or comparison sequence. Cutoff scores which range from 10 to 50 bits are generally used for individual protein families using the SWISS-PROT sequences as model or comparison sequences.

5   Two other algorithms, SIGPEPT and TM, both based on the HMM algorithm described above (see, e.g., Eddy, supra; and Sonnhammer, supra), identify potential signal sequences and transmembrane domains, respectively. SIGPEPT was created using protein sequences having signal sequence annotations derived from SWISS-PROT. It contains about 1413 non-redundant signal sequences ranging in length from 14 to 36 amino acid residues.

10   TM was created similarly using transmembrane domain annotations. It contains about 453 non-redundant transmembrane sequences encompassing 1579 transmembrane domain segments. Suitable HMM models were constructed using the above sequences and were refined with known SWISS-PROT signal peptide sequences or transmembrane domain sequences until a high correlation coefficient, a measurement of the correctness of the

15   analysis, was obtained. Using the protein sequences from the SWISS-PROT database as a test set, a cutoff score of 11 bits, as determined above, correlated with 91-94% true-positives and about 4.1% false-positives, yielding a correlation coefficient of about 0.87-0.90 for SIGPEPT. A score of 11 bits for TM will typically give the following results: 75% true positives; 1.72% false positives; and a correlation coefficient of 0.76. Each search evaluates

20   the statistical significance of any matches found and reports only those matches that score at least 11 bits.

IV.   Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript

25   of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; and Ausubel, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST are used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ™ database (Incyte

30   Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any

particular match is categorized as exact or similar.

The basis of the search is the product score, which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

5 The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

10 The results of Northern analysis are reported as a list of libraries in which the transcript encoding HAPOP occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

15

V. Extension of HAPOP Encoding Polynucleotides

The nucleic acid sequences of Incyte Clones 157658, 642272, 1453807, and 2059022 were used to design oligonucleotide primers for extending partial nucleotide sequences to full length. For each nucleic acid sequence, one primer was synthesized to initiate extension of an

20 antisense polynucleotide, and the other was synthesized to initiate extension of a sense polynucleotide. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO™ 4.06 (National Biosciences, Plymouth, MN), or another appropriate program, to be about 22 to 30

25 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (GIBCO BRL) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to

30 further extend the known region.

53

High fidelity amplification was obtained by following the instructions for the XL-PCR™ kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix. PCR was performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA), beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, with the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat steps 4 through 6 for an additional 15 cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat steps 8 through 10 for an additional 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6% to 0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK™ (QIAGEN Inc.), and trimmed of overhangs using Klenow enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2 to 3 hours, or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) were transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium. (See, e.g., Sambrook, supra, Appendix A, p. 2.) After incubation for one hour at 37°C, the E. coli mixture was plated on Luria Bertani (LB) agar (See, e.g., Sambrook, supra, Appendix A, p. 1) containing carbenicillin (2x carb). The following day, several colonies were randomly picked from each plate and cultured in 150 μl of liquid LB/2x carb medium placed in an individual well of an appropriate commercially-available sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture was transferred into a non-sterile 96-well plate and, after dilution 1:10 with water, 5 μl from

each sample was transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using

5   the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2 through 4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular

15   weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequences of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

20

VI.   Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described,

25   essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO™ 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham, Chicago, IL), and T4 polynucleotide kinase (DuPont NEN®, Boston, MA). The labeled oligonucleotides are substantially purified using a

30   Sephadex™ G-25 superfine size exclusion dextran bead column (Pharmacia & Upjohn, Kalamazoo, MI). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba1, or Pvu II (DuPont NEN,

Boston, MA).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially

5   washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester, NY) is exposed to the blots to film for several hours, hybridization patterns are compared visually.

VII.   Microarrays

10   A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate

15   number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may

20   comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE™. Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using,

25   e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; and Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HAPOP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HAPOP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO™ 4.06 software and the coding sequence of HAPOP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HAPOP-encoding transcript.

IX. Expression of HAPOP

Expression and purification of HAPOP is achieved using bacterial or virus-based expression systems. For expression of HAPOP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HAPOP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HAPOP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant <u>Autographica californica</u> nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HAPOP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect <u>Spodoptera frugiperda</u> (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HAPOP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification

5 of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Pharmacia, Piscataway, NJ). Following purification, the GST moiety can be proteolytically cleaved from HAPOP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak, Rochester, NY). 6-His, a stretch of

10 six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN Inc, Chatsworth, CA). Methods for protein expression and purification are discussed in Ausubel, F. M. et al. (1995 and periodic supplements) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, ch 10, 16. Purified HAPOP obtained by these methods can be used directly in the following activity assay.

15

X.   Demonstration of HAPOP Activity

An assay for HAPOP activity measures the induction of apoptosis when HAPOP is expressed at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high

20 levels of cDNA expression. Vectors of choice include pCMV SPORT™ (Life Technologies, Gaithersburg, MD) and pCR™ 3.1 (Invitrogen, Carlsbad, CA, both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences

25 encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP) (Clontech, Palo Alto, CA), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to

30 identify transfected cells expressing GFP or CD64-GFP and to evaluate their apoptotic state.

FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-

5   regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York, NY.

10

XI. Transcriptional Analysis

The influence of HAPOP on gene expression can also be assessed using highly purified populations of cells transfected with sequences encoding HAPOP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind

15   to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success, NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HAPOP and other genes of interest can be analyzed by Northern analysis or microarray

20   techniques.

XII.   Production of HAPOP Specific Antibodies

HAPOP substantially purified using polyacrylamide gel electrophoresis (PAGE)(see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification

25   techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HAPOP amino acid sequence is analyzed using LASERGENE™ software (DNASTAR Inc.) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the

30   C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel supra,

ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry and coupled to KLH (Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester

5   (MBS) to increase immunogenicity. (See, e.g., Ausubel supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

10

XIII.   Purification of Naturally Occurring HAPOP Using Specific Antibodies

Naturally occurring or recombinant HAPOP is substantially purified by immunoaffinity chromatography using antibodies specific for HAPOP. An immunoaffinity column is constructed by covalently coupling anti-HAPOP antibody to an activated

15   chromatographic resin, such as CNBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HAPOP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HAPOP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under

20   conditions that disrupt antibody/HAPOP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HAPOP is collected.

XIV.   Identification of Molecules Which Interact with HAPOP

HAPOP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter

25   reagent. (See, e.g., Bolton et al. (1973) Biochem. J. 133:529.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HAPOP, washed, and any wells with labeled HAPOP complex are assayed. Data obtained using different concentrations of HAPOP are used to calculate values for the number, affinity, and association of HAPOP with the candidate molecules.

30   Various modifications and variations of the described methods and systems of the

invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described

5 modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Hillman, Jennifer L.
              Corley, Neil C.
              Guegler, Karl J.
              Patterson, Chandra
              Baughn, Mariah

(ii) TITLE OF THE INVENTION: HUMAN APOPTOSIS ASSOCIATED PROTEINS

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
  (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
  (B) STREET: 3174 Porter Dr.
  (C) CITY: Palo Alto
  (D) STATE: CA
  (E) COUNTRY: USA
  (F) ZIP: 94304

(v) COMPUTER READABLE FORM:
  (A) MEDIUM TYPE: Diskette
  (B) COMPUTER: IBM Compatible
  (C) OPERATING SYSTEM: DOS
  (D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
  (A) APPLICATION NUMBER: To Be Assigned
  (B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:
  (A) APPLICATION NUMBER:
  (B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
  (A) NAME: Cerrone, Michael C.
  (B) REGISTRATION NUMBER: 39,132
  (C) REFERENCE/DOCKET NUMBER: PF-0519 US

(ix) TELECOMMUNICATION INFORMATION:
  (A) TELEPHONE: 650-855-0555
  (B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 480 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
  (A) LIBRARY: THP1PLB02
  (B) CLONE: 157658

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ser Ala Glu Val Ile His Gln Val Glu Glu Ala Leu Asp Thr Asp
1               5                   10                  15
Glu Lys Glu Met Leu Leu Phe Leu Cys Arg Asp Val Ala Ile Asp Val
            20                  25                  30
```

```
Val Pro Pro Asn Val Arg Asp Leu Leu Asp Ile Leu Arg Glu Arg Gly
        35              40              45
Lys Leu Ser Val Gly Asp Leu Ala Glu Leu Leu Tyr Arg Val Arg Arg
    50              55              60
Phe Asp Leu Leu Lys Arg Ile Leu Lys Met Asp Arg Lys Ala Val Glu
65              70              75              80
Thr His Leu Leu Arg Asn Pro His Leu Val Ser Asp Tyr Arg Val Leu
            85              90              95
Met Ala Glu Ile Gly Glu Asp Leu Asp Lys Ser Asp Val Ser Ser Leu
            100             105             110
Ile Phe Leu Met Lys Asp Tyr Met Gly Arg Gly Lys Ile Ser Lys Glu
        115             120             125
Lys Ser Phe Leu Asp Leu Val Val Glu Leu Glu Lys Leu Asn Leu Val
    130             135             140
Ala Pro Asp Gln Leu Asp Leu Leu Glu Lys Cys Leu Lys Asn Ile His
145             150             155             160
Arg Ile Asp Leu Lys Thr Lys Ile Gln Lys Tyr Lys Gln Ser Val Gln
            165             170             175
Gly Ala Gly Thr Ser Tyr Arg Asn Val Leu Gln Ala Ala Ile Gln Lys
        180             185             190
Ser Leu Lys Asp Pro Ser Asn Asn Phe Arg Leu His Asn Gly Arg Ser
        195             200             205
Lys Glu Gln Arg Leu Lys Glu Gln Leu Gly Ala Gln Gln Glu Pro Val
    210             215             220
Lys Lys Ser Ile Gln Glu Ser Glu Ala Phe Leu Pro Gln Ser Ile Pro
225             230             235             240
Glu Glu Arg Tyr Lys Met Lys Ser Lys Pro Leu Gly Ile Cys Leu Ile
            245             250             255
Ile Asp Cys Ile Gly Asn Glu Thr Glu Leu Leu Arg Asp Thr Phe Thr
            260             265             270
Ser Leu Gly Tyr Glu Val Gln Lys Phe Leu His Leu Ser Met His Gly
        275             280             285
Ile Ser Gln Ile Leu Gly Gln Phe Ala Cys Met Pro Glu His Arg Asp
        290             295             300
Tyr Asp Ser Phe Val Cys Val Leu Val Ser Arg Gly Gly Ser Gln Ser
305             310             315             320
Val Tyr Gly Val Asp Gln Thr His Ser Gly Leu Pro Leu His His Ile
            325             330             335
Arg Arg Met Phe Met Gly Asp Ser Cys Pro Tyr Leu Ala Gly Lys Pro
            340             345             350
Lys Met Phe Phe Ile Gln Asn Tyr Val Val Ser Glu Gly Gln Leu Glu
        355             360             365
Asp Ser Ser Leu Leu Glu Val Asp Gly Pro Ala Met Lys Asn Val Glu
    370             375             380
Phe Lys Ala Gln Lys Arg Gly Leu Cys Thr Val His Arg Glu Ala Asp
385             390             395             400
Phe Phe Trp Ser Leu Cys Thr Ala Asp Met Ser Leu Leu Glu Gln Ser
            405             410             415
His Ser Ser Pro Ser Leu Tyr Leu Gln Cys Leu Ser Gln Lys Leu Arg
            420             425             430
Gln Glu Arg Lys Arg Pro Leu Leu Asp Leu His Ile Glu Leu Asn Gly
        435             440             445
Tyr Met Tyr Asp Trp Asn Ser Arg Val Ser Ala Lys Glu Lys Tyr Tyr
    450             455             460
Val Trp Leu Gln His Thr Leu Arg Lys Lys Leu Ile Leu Ser Tyr Thr
465             470             475             480
```

(2) INFORMATION FOR SEQ ID NO:2:

   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 2352 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (vii) IMMEDIATE SOURCE:
       (A) LIBRARY: THP1PLB02

(B) CLONE: 157658

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
GAAATTGCGC CACTGCACTC CAGCCTGGGC CACAGAGCGA GACTCTGTCT CAAAAAAGAA    60
GGAAAGAAAG AAAGAAAAAA AAAAACACTC GCAGTGTTTA CTCCTAACGC GTGGAACTTG   120
TGTCGACATC CACCCCCGGT TACTGCATAC TCAGTCACAC AAGCCATAGC AGGAAACAGC   180
GAGCTTGCAG CCTCACCGAC GAGTCTCAAC TAAAAGGGAC TCCCGGAGCT AGGGGTGGGG   240
ACTCGGCCTC ACACAGTGAG TGCCGGCTAT TGGACTTTTG TCCAGTGACA GCTGAGACAA   300
CAAGGACCAC GGGAGGAGGT GTAGGAGAGA AGCGCCGCGA ACAGCGATCG CCCAGCACCA   360
AGTCCGCTTC CAGGCTTTCG GTTTCTTTGC CTCCATCTTG GGTGCGCCTT CCCGGCGTCT   420
AGGGGAGCGA AGGCTGAGGT GGCAGCGGCA GGAGAGTCCG GCCGCGACAG GACGAACTCC   480
CCCACTGGAA AGGATTCTGA AAGAAATGAA GTCAGCCCTC AGAAATGAAG TTGACTGCCT   540
GCTGGCTTTC TGTTGACTGG CCCGGAGCTG TACTGCAAGA CCCTTGTGAG CTTCCCTAGT   600
CTAAGAGTAG GATGTCTGCT GAAGTCATCC ATCAGGTTGA AGAAGCACTT GATACAGATG   660
AGAAGGAGAT GCTGCTCTTT TTGTGCCGGG ATGTTGCTAT AGATGTGGTT CCACCTAATG   720
TCAGGGACCT TCTGGATATT TTACGGGAAA GAGGTAAGCT GTCTGTCGGG GACTTGGCTG   780
AACTGCTCTA CAGAGTGAGG CGATTTGACC TGCTCAAACG TATCTTGAAG ATGGACAGAA   840
AAGCTGTGGA GACCCACCTG CTCAGGAACC CTCACCTTGT TTCGGACTAT AGAGTGCTGA   900
TGGCAGAGAT TGGTGAGGAT TTGGATAAAT CTGATGTGTC CTCATTAATT TTCCTCATGA   960
AGGATTACAT GGGCCGAGGC AAGATAAGCA AGGAGAAGAG TTTCTTGGAC CTTGTGGTTG  1020
AGTTGGAGAA ACTAAATCTG GTTGCCCCAG ATCAACTGGA TTTATTAGAA AAATGCCTAA  1080
AGAACATCCA CAGAATAGAC CTGAAGACAA AAATCCAGAA GTACAAGCAG TCTGTTCAAG  1140
GAGCAGGGAC AAGTTACAGG AATGTTCTCC AAGCAGCAAT CCAAAAGAGT CTCAAGGATC  1200
CTTCAAATAA CTTCAGGCTC CATAATGGGA GAAGTAAAGA ACAAAGACTT AAGGAACAGC  1260
TTGGCGCTCA ACAAGAACCA GTGAAGAAAT CCATTCAGGA ATCAGAAGCT TTTTTGCCTC  1320
AGAGCATACC TGAAGAGAGA TACAAGATGA AGAGCAAGCC CCTAGGAATC TGCCTGATAA  1380
TCGATTGCAT TGGCAATGAG ACAGAGCTTC TTCGAGACAC CTTCACTTCC CTGGGCTATG  1440
AAGTCCAGAA ATTCTTGCAT CTCAGTATGC ATGGTATATC CCAGATTCTT GGCCAATTTG  1500
CCTGTATGCC CGAGCACCGA GACTACGACA GCTTTGTGTG TGTCCTGGTG AGCCGAGGAG  1560
GCTCCCAGAG TGTGTATGGT GTGGATCAGA CTCACTCCGG GCTCCCCCTG CATCACATCA  1620
GGAGGATGTT CATGGGAGAT TCATGCCCTT ATCTAGCAGG GAAGCCAAAG ATGTTTTTTA  1680
TTCAGAACTA TGTGGTGTCA GAGGGCCAGC TGGAGGACAG CAGCCTCTTG GAGGTGGATG  1740
GGCCAGCGAT GAAGAATGTG GAATTCAAGG CTCAGAAGCG AGGGCTGTGC ACAGTTCACC  1800
GAGAAGCTGA CTTCTTCTGG AGCCTGTGTA CTGCGGACAT GTCCCTGCTG GAGCAGTCTC  1860
ACAGCTCACC ATCCCTGTAC CTGCAGTGCC TCTCCCAGAA ACTGAGACAA GAAAGAAAAC  1920
GCCCACTCCT GGATCTTCAC ATTGAACTCA ATGGCTACAT GTATGATTGG AACAGCAGAG  1980
TTTCTGCCAA GGAGAAATAT TATGTCTGGC TGCAGCACAC TCTGAGAAAG AAACTTATCC  2040
TCTCCTACAC ATAAGAAACC AAAAGGCTGG GCGTAGTGGC TCACACCTGT GATCCCAGCA  2100
CTTTGGGAGG CCAAGGAGGG CAGATCACTT CAGGTCAGGA GTTCGAGACC AGCCTGGCCA  2160
ACATGGTAAA CGCTGTCCCT AGTAAAAATA CAAAAATTAG CTGGGTGTGG GTGTGGGTAC  2220
CTGTATTCCC AGTTACTTGG GAGGCTGAGG TGGGAGGATC TTTTGAACCC AGGAGTTCAG  2280
GGTCATAGCA TGCTGTGATT GTGCCTACGA ATAGCCACTG CATACCAACC TGGGCAATAT  2340
AGCAAGATCC CA                                                     2352
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 238 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:
    (A) LIBRARY: BRSTNOT03
    (B) CLONE: 642272

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Glu Tyr Ala Met Lys Ser Leu Ser Leu Leu Tyr Pro Lys Ser Leu
1               5                   10                  15
Ser Arg His Val Ser Val Arg Thr Ser Val Val Thr Gln Gln Leu Leu
            20                  25                  30
Ser Glu Pro Ser Pro Lys Ala Pro Arg Ala Arg Pro Cys Arg Val Ser
        35                  40                  45
Thr Ala Asp Arg Ser Val Arg Lys Gly Ile Met Ala Tyr Ser Leu Glu
    50                  55                  60
```

64

```
Asp Leu Leu Leu Lys Val Arg Asp Thr Leu Met Leu Ala Asp Lys Pro
65               70              75                  80
Phe Phe Leu Val Leu Glu Glu Asp Gly Thr Thr Val Glu Thr Glu Glu
            85              90              95
Tyr Phe Gln Ala Leu Ala Gly Asp Thr Val Phe Met Val Leu Gln Lys
        100             105                 110
Gly Gln Lys Trp Gln Pro Pro Ser Glu Gln Gly Thr Arg His Pro Leu
        115             120                 125
Ser Leu Ser His Lys Pro Ala Lys Lys Ile Asp Val Ala Arg Val Thr
    130             135             140
Phe Asp Leu Tyr Lys Leu Asn Pro Gln Asp Phe Ile Gly Cys Leu Asn
145             150                 155                 160
Val Lys Ala Thr Phe Tyr Asp Thr Tyr Ser Leu Ser Tyr Asp Leu His
            165             170                 175
Cys Cys Gly Ala Lys Arg Ile Met Lys Glu Ala Phe Arg Trp Ala Leu
        180             185                 190
Phe Ser Met Gln Ala Thr Gly His Val Leu Leu Gly Thr Ser Cys Tyr
    195             200                 205
Leu Gln Gln Leu Leu Asp Ala Thr Glu Glu Gly Gln Pro Pro Lys Gly
    210             215                 220
Lys Ala Ser Ser Leu Ile Pro Thr Cys Leu Lys Ile Leu Gln
225             230             235
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 1284 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

    (vii) IMMEDIATE SOURCE:
      (A) LIBRARY: BRSTNOT03
      (B) CLONE: 642272

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
AATGTTCTTT TGGCCACTGT GAAGCCTCAG GAAGGGGCTC GGATTGCTCA AGGACCCATG      60
GGAGAGAGGA GGCTTTGACT GGGCTGCCTG CCTGTGAGGT CTCTGGACTA GAGGTCCAAC     120
GCAGTCCAGC TGACAAGGAT GGAATACGCC ATGAAGTCCC TTAGCCTTCT CTACCCCAAG     180
TCCCTCTCCA GGCATGTGTC AGTGCGTACC TCTGTGGTGA CCCAGCAGCT GCTGTCGGAG     240
CCCAGCCCCA AGGCCCCCAG GGCCCGGCCC TGCCGCGTAA GCACGGCGGA TCGAAGCGTG     300
AGGAAGGGCA TCATGGCTTA CAGTCTTGAG GACCTCCTCC TCAAGGTCCG GGACACTCTG     360
ATGCTGGCAG ACAAGCCCTT CTTCCTGGTG CTGGAGGAAG ATGGCACAAC TGTAGAGACA     420
GAAGAGTACT TCCAAGCCCT GGCAGGGGAT ACAGTGTTCA TGGTCCTCCA GAAGGGGCAG     480
AAATGGCAGC CCCCATCAGA ACAGGGGACA AGGCACCCAC TGTCCCTCTC CCATAAGCCT     540
GCCAAGAAGA TTGATGTGGC CCGTGTAACG TTTGATCTGT ACAAGCTGAA CCCACAGGAC     600
TTCATTGGCT GCCTGAACGT GAAGGCGACT TTTTATGATA CATACTCCCT TTCCTATGAT     660
CTGCACTGCT GTGGGGCCAA GCGCATCATG AAGGAAGCTT TCCGCTGGGC CCTCTTCAGC     720
ATGCAGGCCA CAGGCCACGT ACTGCTTGGC ACCTCCTGTT ACCTGCAGCA GCTCCTCGAT     780
GCTACGGAGG AAGGGCAGCC CCCCAAGGGC AAGGCCTCAT CCCTTATCCC GACCTGTCTG     840
AAGATACTGC AGTGAAAGCC CAAGTCCTTG GAAGCTTTCC CCAGTGAAGG ACTGACTGGG     900
GGCCTCACGC TTAACTGGTA GTGCCCACAA GCCTGGCAGC TGTAGAGCCG CGAACCTCCC     960
CACACCTCCC TCACCGCGCA GGACCCTGAC TGAGGAGGAG GAGCTGGAAA CCTGGGGTGG    1020
GTTGGCCAAA GGAGAACCTC AAGCTCCTGG CCTGATCCAG CTCCTTCCTG CCCAAGGCAG    1080
CTTAGCCCAT CCAGACTGGT CCTGAAGTCT GTCCCTCCAT TGGCATGAAG TCTGCCCCTT    1140
AGCAATCCGG CCTCGCAGGC TGTACTTTCA TGGTGCTCTC TACCTTCTGG CCCCCATCCC    1200
GGAACATTCC TGAGTGAATT CGCAAGCGCA CTAGCATGTG ATATTAGGGA GTTTGCAATA    1260
AATTATTGAG GCTGAAAAAA AAAA                                          1284
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 410 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

```
        (vii)  IMMEDIATE SOURCE:
               (A)  LIBRARY: PENITUT01
               (B)  CLONE: 1453807

        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Leu Gln Glu Gln Val Ser Glu Tyr Leu Gly Val Thr Ser Phe Lys
  1               5                  10                  15
Arg Lys Tyr Pro Asp Leu Glu Arg Arg Asp Leu Ser His Lys Glu Lys
             20                  25                  30
Leu Tyr Leu Arg Glu Leu Asn Val Ile Thr Glu Thr Gln Cys Thr Leu
         35                  40                  45
Gly Leu Thr Ala Leu Arg Ser Asp Glu Val Ile Asp Leu Met Ile Lys
     50                  55                  60
Glu Tyr Pro Ala Lys His Ala Glu Tyr Ser Val Ile Leu Gln Glu Lys
65                  70                  75                  80
Glu Arg Gln Arg Ile Thr Asp His Tyr Lys Glu Tyr Ser Gln Met Gln
                 85                  90                  95
Gln Gln Asn Thr Gln Lys Val Glu Ala Ser Lys Val Pro Glu Tyr Ile
            100                 105                 110
Lys Lys Ala Ala Lys Lys Ala Ala Glu Phe Asn Ser Asn Leu Asn Arg
        115                 120                 125
Glu Arg Met Glu Glu Arg Arg Ala Tyr Phe Asp Leu Gln Thr His Val
        130                 135                 140
Ile Gln Val Pro Gln Gly Lys Tyr Lys Val Leu Pro Thr Glu Arg Thr
145                 150                 155                 160
Lys Val Ser Ser Tyr Pro Val Ala Leu Ile Pro Gly Gln Phe Gln Glu
                165                 170                 175
Tyr Tyr Lys Arg Tyr Ser Pro Asp Glu Leu Arg Tyr Leu Pro Leu Asn
            180                 185                 190
Thr Ala Leu Tyr Glu Pro Pro Leu Asp Pro Glu Leu Pro Ala Leu Asp
        195                 200                 205
Ser Asp Gly Asp Ser Asp Asp Gly Glu Asp Gly Arg Gly Asp Glu Lys
        210                 215                 220
Arg Lys Asn Lys Gly Thr Ser Asp Ser Ser Ser Gly Asn Val Ser Glu
225                 230                 235                 240
Gly Glu Ser Pro Pro Asp Ser Gln Glu Asp Ser Phe Gln Gly Arg Gln
                245                 250                 255
Lys Ser Lys Asp Lys Ala Ala Thr Pro Arg Lys Asp Gly Pro Lys Arg
            260                 265                 270
Ser Val Leu Ser Lys Ser Val Pro Gly Tyr Lys Pro Lys Val Ile Pro
        275                 280                 285
Asn Ala Ile Cys Gly Ile Cys Leu Lys Gly Lys Glu Ser Asn Lys Lys
        290                 295                 300
Gly Lys Ala Glu Ser Leu Ile His Cys Ser Gln Cys Glu Asn Ser Gly
305                 310                 315                 320
His Pro Ser Cys Leu Asp Met Thr Met Glu Leu Val Ser Met Ile Lys
                325                 330                 335
Thr Tyr Pro Trp Gln Cys Met Glu Cys Lys Thr Cys Ile Ile Cys Gly
            340                 345                 350
Gln Pro His His Glu Glu Glu Met Met Phe Cys Asp Met Cys Asp Arg
        355                 360                 365
Gly Tyr His Thr Phe Cys Val Gly Leu Gly Ala Ile Pro Ser Gly Arg
        370                 375                 380
Trp Ile Cys Asp Cys Cys Gln Arg Ala Pro Pro Thr Pro Arg Lys Val
385                 390                 395                 400
Gly Arg Arg Gly Lys Asn Ser Lys Glu Gly
            405                 410

        (2)  INFORMATION FOR SEQ ID NO:6:

        (i)  SEQUENCE CHARACTERISTICS:
             (A)  LENGTH: 2150 base pairs
             (B)  TYPE: nucleic acid
             (C)  STRANDEDNESS: single
             (D)  TOPOLOGY: linear
```

(vii) IMMEDIATE SOURCE:
   (A) LIBRARY: PENITUT01
   (B) CLONE: 1453807

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
CTTGGTTTTA GTTACTATCC AGCAGAAAAC TTGATAGAGT ACAAATGGCC ACCTGATGAA     60
ACAGGAGAAT ACTATATGCT TCAAGAACAA GTCAGTGAAT ATTTGGGTGT GACCTCCTTT    120
AAAAGGAAAT ATCCAGATTT AGAGCGACGA GATTTGTCTC ACAAGGAGAA ACTCTACCTG    180
AGAGAGCTAA ATGTCATTAC TGAAACTCAG TGCACTCTAG GCTTAACAGC ATTGCGCAGT    240
GATGAAGTGA TTGATTTAAT GATAAAAGAA TATCCAGCCA AACATGCTGA GTATTCTGTT    300
ATTCTACAAG AAAAAGAACG TCAACGAATT ACAGACCATT ATAAAGAGTA TTCCCAAATG    360
CAACAACAGA ATACTCAGAA AGTTGAAGCC AGTAAAGTGC CTGAGTATAT TAAGAAAGCT    420
GCCAAAAAAG CAGCAGAATT TAATAGCAAC TTAAACCGGG AACGCATGGA AGAAAGAAGA    480
GCTTATTTTG ACTTGCAGAC ACATGTTATC CAGGTACCTC AAGGGAAGTA CAAAGTTTTG    540
CCAACAGAGC GAACAAAGGT CAGTTCTTAC CCAGTGGCTC TCATCCCCGG ACAGTTCCAG    600
GAATATTATA AGAGGTACTC ACCAGATGAG CTGCGGTATC TGCCATTAAA CACAGCCCTG    660
TATGAGCCCC CTCTGGATCC TGAGCTCCCT GCTCTAGACA GTGATGGTGA TTCAGATGAT    720
GGCGAAGATG GTCGAGGTGA TGAGAAACGG AAAAATAAAG GCACTTCGGA CAGCTCCTCT    780
GGCAATGTAT CTGAAGGGGA AAGCCCTCCT GACAGCCAGG AGGACTCTTT CCAGGGAAGA    840
CAGAAATCAA AAGACAAAGC TGCCACTCCA AGAAAGATG GTCCCAAACG TTCTGTACTG    900
TCCAAGTCAG TTCCTGGGTA CAAGCCAAAG GTCATTCCAA ATGCTATATG TGGAATTTGT    960
CTGAAGGGTA AGGAGTCCAA CAAGAAAGGA AAGGCTGAAT CACTTATACA CTGCTCCCAA   1020
TGTGAGAATA GTGGCCATCC TTCTTGCCTG GATATGACAA TGGAGCTTGT TTCTATGATT   1080
AAGACCTACC CATGGCAGTG TATGGAATGT AAAACATGCA TTATATGTGG ACAACCCCAC   1140
CATGAAGAAG AAATGATGTT CTGTGATATG TGTGACAGAG GTTATCATAC TTTTTGTGTG   1200
GGCCTTGGTG CTATTCCATC AGGTCGCTGG ATTTGTGACT GTTGTCAGCG GGCCCCCCCA   1260
ACACCCAGGA AAGTGGGCAG AAGGGGGAAA AACAGCAAAG AGGGATAAAA TAGTTTTTGA   1320
CTCTAATACT GTATATGCAT TTAAGTGGAA TATTTGGTGC CATTTACAAC ATTATTTTCA   1380
TGCCAATAAA AGATTTTTTT TGCAAATTAT GAGCTTAAAA TCTGCAGTTA TTTCTGTTAA   1440
AAGTACGCTT ACTCTCGAAA CTAACTCCAG GTAGAGAATT CATCTTCCAA AGTATTTTAT   1500
AGTAACCTTG GCTCACTCCA AAAATTCAGT GGAAATGTTT AGTAACTTAA GATACTTAAC   1560
TGTTTCTCCA TAGCCCCAAA AGTTAATTTT CATGAAACTT CCTAATCTAC ATTGTTTCCG   1620
GCCTACCATA GGTAGCACTG ACAAAGTTAT TTAATAACTG AAGAATTTTC ATAGGTATGA   1680
CAAATGGCCC ACTAAGATTT GGTGCAGCTG GATTTAGAGT TGTCATTATT GGACTGGTAC   1740
AGGAACAAAC TTTGTAAATA CCTGCCTGCC AGGAAATCCT TTTTGTATAG AAAAGTACCA   1800
TCACCTACTT GGGGTACAGG CATGAGGCTT TAGTCCAGGC TCAGGGAAGT GTACGTAAAT   1860
CATTTCCAAC TTGATTTTAG TAACTCTTGA AAACTTACAC CAACTTCGGT TAGAATCTCC   1920
AGAGTAAAAT TACAAAGTTA TCAACCTTTT GATTTGTGTC ACAGCATGAA AGGTTGCTCT   1980
ATTTTATATA AACCTGTTAC TGCAATCATT TTTAGTCAAC CTGCCTAATG AAAATGGAGT   2040
CTAAACACTT TGTGCACAGT CCTTTTATAG GAATTATGAT CTTTAAAATA CTGTGCTTGC   2100
TGCTTTTCCT ATTTTTGGGG TAACTGAGGT AACAAAATGC GTATGGCTTT              2150
```

(2) INFORMATION FOR SEQ ID NO:7:

   (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 211 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

   (vii) IMMEDIATE SOURCE:
      (A) LIBRARY: OVARNOT03
      (B) CLONE: 2059022

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Ala Asn Ser Gly Leu Gln Leu Leu Gly Phe Ser Met Ala Leu Leu
1               5                   10                  15
Gly Trp Val Gly Leu Val Ala Cys Thr Ala Ile Pro Gln Trp Gln Met
            20                  25                  30
Ser Ser Tyr Ala Gly Asp Asn Ile Ile Thr Ala Gln Ala Met Tyr Lys
        35                  40                  45
Gly Leu Trp Met Asp Cys Val Thr Gln Ser Thr Gly Met Met Ser Cys
    50                  55                  60
Lys Met Tyr Asp Ser Val Leu Ala Leu Ser Ala Ala Leu Gln Ala Thr
65                  70                  75                  80
```

```
Arg Ala Leu Met Val Val Ser Leu Val Leu Gly Phe Leu Ala Met Phe
                85              90              95
Val Ala Thr Met Gly Met Lys Cys Thr Arg Cys Gly Gly Asp Asp Lys
            100             105             110
Val Lys Lys Ala Arg Ile Ala Met Gly Gly Gly Ile Ile Phe Ile Val
        115             120             125
Ala Gly Leu Ala Ala Leu Val Ala Cys Ser Trp Tyr Gly His Gln Ile
    130             135             140
Val Thr Asp Phe Tyr Asn Pro Leu Ile Pro Thr Asn Ile Lys Tyr Glu
145             150             155             160
Phe Gly Pro Ala Ile Phe Ile Gly Trp Ala Gly Ser Ala Leu Val Ile
            165             170             175
Leu Gly Gly Ala Leu Leu Ser Cys Ser Cys Pro Gly Asn Glu Ser Lys
            180             185             190
Ala Gly Tyr Arg Ala Pro Arg Ser Tyr Pro Lys Ser Asn Ser Ser Lys
    195             200             205
Glu Tyr Val
    210
```

        (2) INFORMATION FOR SEQ ID NO:8:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 1546 base pairs
                (B) TYPE: nucleic acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

            (vii) IMMEDIATE SOURCE:
                (A) LIBRARY: OVARNOT03
                (B) CLONE: 2059022

            (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
CCCACGCGTC CGCTCACCTC CGAGCCACCT CTGCTGCGCA CCGCAGCCTC GGACCTACAG      60
CCCAGGATAC TTTGGGACTT GCCGGCGCTC AGAAACGCGC CCAGACGGCC CCTCCACCTT     120
TTGTTTGCCT AGGGTCGCCG AGAGCGCCCG GAGGGAACCG CCTGGCCTTC GGGGACCACC     180
AATTTTGTCT GGAACCACCC TCCCGGCGTA TCCTACTCCC TGTGCCGCGA GGCCATCGCT     240
TCACTGGAGG GGTCGATTTG TGTGTAGTTT GGTGACAAGA TTTGCATTCA CCTGGCCCAA     300
ACCCTTTTTG TCTCTTTGGG TGACCGGAAA ACTCCACCTC AAGTTTTCTT TTGTGGGGCT     360
GCCCCCCAAG TGTCGTTTGT TTTACTGTAG GGTCTCCCCG CCCGGCGCCC CCAGTGTTTT     420
CTGAGGGCGG AAATGGCCAA TTCGGGCCTG CAGTTGCTGG GCTTCTCCAT GGCCCTGCTG     480
GGCTGGGTGG GTCTGGTGGC CTGCACCGCC ATCCCGCAGT GGCAGATGAG CTCCTATGCG     540
GGTGACAACA TCATCACGGC CCAGGCCATG TACAAGGGGC TGTGGATGGA CTGCGTCACG     600
CAGAGCACGG GGATGATGAG CTGCAAAATG TACGACTCGG TGCTCGCCCT GTCCGCGGCC     660
TTGCAGGCCA CTCGAGCCCT AATGGTGGTC TCCCTGGTGC TGGGCTTCCT GGCCATGTTT     720
GTGGCCACGA TGGGCATGAA GTGCACGCGC TGTGGGGGAG ACGACAAAGT GAAGAAGGCC     780
CGTATAGCCA TGGGTGGAGG CATAATTTTC ATCGTGGCAG GTCTTGCCGC CTTGGTAGCT     840
TGCTCCTGGT ATGGCCATCA GATTGTCACA GACTTTTATA ACCCTTTGAT CCCTACCAAC     900
ATTAAGTATG AGTTTGGCCC TGCCATCTTT ATTGGCTGGG CAGGGTCTGC CCTAGTCATC     960
CTGGGAGGTG CACTGCTCTC CTGTTCCTGT CCTGGGAATG AGAGCAAGGC TGGGTACCGT    1020
GCACCCCGCT CTTACCCTAA GTCCAACTCT TCCAAGGAGT ATGTGTGACC TGGGATCTCC    1080
TTGCCCCAGC CTGACAGGCT ATGGGAGTGT CTAGATGCCT GAAAGGGCCT GGGGCTGAGC    1140
TCAGCCTGTG GGCAGGGTGC CGGACAAAGG CCTCCTGGTC ACTCTGTCCC TGCACTCCAT    1200
GTATAGTCCT CTTGGGTTGG GGGTGGGGGG GTGCCGTTGG TGGGAGAGAC AAAAAGAGGG    1260
AGAGTGTGCT TTTTGTACAG TAATAAAAAA TAAGTATTGG GAAGCAGGCT TTTTTCCCTT    1320
CAGGGCCTCT GCTTTCCTCC CGTCCAGATC CTTGCAGGGA GCTTGGAACC TTAGTGCACC    1380
TACTTCAGTT CAGAACACTT AGCACCCCAC TGACTCCACT GACAATTGAC TAAAAGATGC    1440
AGGTGCTCGT ATCTCGACAT TCATTCCCAC CCCCCTCTTA TTTAAATAGC TACCAAAGTA    1500
CTTCTTTTTT AATAAAAAAA TAAAGATTTT TATTAAAAAA AAAAAA               1546
```

        (2) INFORMATION FOR SEQ ID NO:9:

            (i) SEQUENCE CHARACTERISTICS:
                (A) LENGTH: 239 amino acids
                (B) TYPE: amino acid
                (C) STRANDEDNESS: single
                (D) TOPOLOGY: linear

```
          (vii)  IMMEDIATE SOURCE:
                (A)  LIBRARY: Swiss-Prot
                (B)  CLONE: p56198

          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:9:
```

```
Met Asp Tyr Ala Met Lys Ser Leu Ser Leu Leu Tyr Pro Arg Ser Leu
 1               5                  10                  15
Ser Arg His Val Ala Val Ser Thr Ala Val Val Thr Gln Gln Leu Val
            20                  25                  30
Ser Lys Pro Ser Arg Glu Thr Pro Arg Ala Arg Pro Cys Arg Val Ser
        35                  40                  45
Thr Ala Asp Arg Lys Val Arg Lys Gly Ile Met Ala His Ser Leu Glu
    50                  55                  60
Asp Leu Leu Asn Lys Val Gln Asp Ile Leu Lys Leu Lys Asp Lys Pro
65                  70                  75                  80
Phe Ser Leu Val Leu Glu Glu Asp Gly Thr Ile Val Glu Thr Glu Glu
                85                  90                  95
Tyr Phe Gln Ala Leu Ala Lys Asp Thr Met Phe Met Val Leu Leu Lys
            100                 105                 110
Gly Gln Lys Trp Lys Pro Pro Ser Glu Gln Arg Lys Lys Arg Ala Gln
        115                 120                 125
Leu Ala Leu Ser Gln Lys Pro Thr Lys Lys Ile Asp Val Ala Arg Val
    130                 135                 140
Thr Phe Asp Leu Tyr Lys Leu Asn Pro Gln Asp Phe Ile Gly Cys Leu
145                 150                 155                 160
Asn Val Lys Ala Thr Leu Tyr Asp Thr Tyr Ser Leu Ser Tyr Asp Leu
            165                 170                 175
His Cys Tyr Lys Ala Lys Arg Ile Val Lys Glu Ile Val Arg Trp Thr
            180                 185                 190
Leu Phe Ser Met Gln Ala Thr Gly His Met Leu Leu Gly Thr Ser Ser
            195                 200                 205
Tyr Met Gln Gln Phe Leu Asp Ala Thr Glu Glu Gln Pro Ala Lys
    210                 215                 220
Ala Lys Pro Ser Ser Leu Leu Pro Ala Cys Leu Lys Met Leu Gln
225                 230                 235
```

```
     (2)  INFORMATION FOR SEQ ID NO:10:

          (i)  SEQUENCE CHARACTERISTICS:
                (A)  LENGTH: 331 amino acids
                (B)  TYPE: amino acid
                (C)  STRANDEDNESS: single
                (D)  TOPOLOGY: linear

          (vii)  IMMEDIATE SOURCE:
                (A)  LIBRARY: GenBank
                (B)  CLONE: 2065561

          (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:10:
```

```
Met Glu Val Thr Gly Asp Ala Gly Val Pro Glu Ser Gly Glu Ile Arg
 1               5                  10                  15
Thr Leu Lys Pro Cys Leu Leu Arg Arg Asn Tyr Ser Arg Glu Gln His
            20                  25                  30
Gly Val Ala Ala Ser Cys Leu Glu Asp Leu Arg Ser Lys Ala Cys Asp
        35                  40                  45
Ile Leu Ala Ile Asp Lys Ser Leu Thr Pro Val Thr Leu Val Leu Ala
    50                  55                  60
Glu Asp Gly Thr Ile Val Asp Asp Asp Tyr Phe Leu Cys Leu Pro
65                  70                  75                  80
Ser Asn Thr Lys Phe Val Ala Leu Ala Ser Asn Glu Lys Trp Ala Tyr
                85                  90                  95
Asn Asn Ser Asp Gly Gly Thr Ala Trp Ile Ser Gln Glu Ser Phe Asp
            100                 105                 110
Val Asp Glu Thr Asp Ser Gly Ala Gly Leu Lys Trp Lys Asn Val Ala
        115                 120                 125
```

Arg Gln Leu Lys Glu Asp Leu Ser Ser Ile Ile Leu Leu Ser Glu Glu
    130   135   140

Asp Leu Gln Met Leu Val Asp Ala Pro Cys Ser Asp Leu Ala Gln Glu
145   150   155   160

Leu Arg Gln Ser Cys Ala Thr Val Gln Arg Leu Gln His Thr Leu Gln
              165   170   175

Gln Val Leu Asp Gln Arg Glu Glu Val Arg Gln Ser Lys Gln Leu Leu
        180   185   190

Gln Leu Tyr Leu Gln Ala Leu Glu Lys Glu Gly Ser Leu Leu Ser Lys
      195   200   205

Gln Glu Glu Ser Lys Ala Ala Phe Gly Glu Glu Val Asp Ala Val Asp
    210   215   220

Thr Gly Ile Ser Arg Glu Thr Ser Ser Asp Val Ala Leu Ala Ser His
225   230   235   240

Ile Leu Thr Ala Leu Arg Glu Lys Gln Ala Pro Glu Leu Ser Leu Ser
              245   250   255

Ser Gln Asp Leu Glu Leu Val Thr Lys Glu Asp Pro Lys Ala Leu Ala
        260   265   270

Val Ala Leu Asn Trp Asp Ile Lys Lys Thr Glu Thr Val Gln Glu Ala
      275   280   285

Cys Glu Arg Glu Leu Ala Leu Arg Leu Gln Gln Thr Gln Ser Leu His
      290   295   300

Ser Leu Arg Ser Ile Ser Ala Ser Lys Ala Ser Pro Pro Gly Asp Leu
305   310   315   320

Gln Asn Pro Lys Arg Ala Arg Gln Asp Pro Thr
        325   330

       (2) INFORMATION FOR SEQ ID NO:11:

        (i) SEQUENCE CHARACTERISTICS:
           (A) LENGTH: 371 amino acids
           (B) TYPE: amino acid
           (C) STRANDEDNESS: single
           (D) TOPOLOGY: linear

      (vii) IMMEDIATE SOURCE:
           (A) LIBRARY: GenBank
           (B) CLONE: 606661

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Glu Gln Cys His Asn Tyr Asn Ala Arg Leu Cys Ala Glu Arg Ser
1   5   10   15

Val Arg Leu Pro Phe Leu Asp Ser Gln Thr Gly Val Ala Gln Ser Asn
        20   25   30

Cys Tyr Ile Trp Met Glu Lys Arg His Arg Gly Pro Gly Leu Ala Ser
        35   40   45

Gly Gln Leu Tyr Ser Tyr Pro Ala Arg Arg Trp Arg Lys Lys Arg Arg
    50   55   60

Ala His Pro Pro Glu Asp Pro Arg Leu Ser Phe Pro Ser Ile Lys Pro
65   70   75   80

Asp Thr Asp Gln Thr Leu Lys Lys Glu Gly Leu Ile Ser Gln Asp Gly
        85   90   95

Ser Ser Leu Glu Ala Leu Leu Arg Thr Asp Pro Leu Glu Lys Arg Gly
        100   105   110

Ala Pro Asp Pro Arg Val Asp Asp Asp Ser Leu Gly Glu Phe Pro Val
        115   120   125

Ser Asn Ser Arg Ala Arg Lys Arg Ile Ile Glu Pro Asp Asp Phe Leu
        130   135   140

Asp Asp Leu Asp Asp Glu Asp Tyr Glu Glu Asp Arg Pro Lys Arg Arg
·145   150   155   160

Gly Lys Gly Lys Ser Lys Ser Lys Gly Val Ser Ser Ala Arg Lys Lys
              165   170   175

Leu Asp Ala Ser Ile Leu Glu Asp Arg Asp Lys Pro Tyr Ala Cys Asp
        180   185   190

Ile Cys Gly Lys Arg Tyr Lys Asn Arg Pro Gly Leu Ser Tyr His Tyr
        195   200   205

```
Ala His Ser His Leu Ala Glu Glu Gly Glu Asp Lys Glu Asp Ser
    210             215             220
Arg Pro Pro Thr Pro Val Ser Gln Arg Ser Glu Glu Gln Lys Ser Lys
225             230             235             240
Lys Gly Pro Asp Gly Leu Ala Leu Pro Asn Asn Tyr Cys Asp Phe Cys
            245             250             255
Leu Gly Asp Ser Lys Ile Asn Lys Lys Thr Gly Gln Pro Glu Glu Leu
            260             265             270
Val Ser Cys Ser Asp Cys Gly Arg Ser Gly His Pro Ser Cys Leu Gln
            275             280             285
Phe Thr Pro Val Met Met Ala Ala Val Lys Thr Tyr Arg Trp Gln Cys
    290             295             300
Ile Glu Cys Lys Cys Cys Asn Leu Cys Gly Thr Ser Glu Asn Asp Asp
305             310             315             320
Gln Leu Leu Phe Cys Asp Asp Cys Asp Arg Gly Tyr His Met Tyr Cys
            325             330             335
Leu Thr Pro Ser Met Ser Glu Pro Pro Glu Gly Ser Trp Ser Cys His
            340             345             350
Leu Cys Leu Asp Leu Leu Lys Glu Lys Ala Ser Ile Tyr Gln Asn Gln
    355             360             365
Asn Ser Ser
370
```

        (2) INFORMATION FOR SEQ ID NO:12:

        (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 220 amino acids
            (B) TYPE: amino acid
            (C) STRANDEDNESS: single
            (D) TOPOLOGY: linear

        (vii) IMMEDIATE SOURCE:
            (A) LIBRARY: GenBank
            (B) CLONE: 2570129

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Ser Met Gly Leu Glu Ile Thr Gly Thr Ala Leu Ala Val Leu Gly
  1             5             10             15
Trp Leu Gly Thr Ile Val Cys Cys Ala Leu Pro Met Trp Arg Val Ser
            20             25             30
Ala Phe Ile Gly Ser Asn Ile Ile Thr Ser Gln Asn Ile Trp Glu Gly
            35             40             45
Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly Gln Met Gln Cys Lys
    50             55             60
Val Tyr Asp Ser Leu Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala Arg
65             70             75             80
Ala Leu Ile Val Val Ala Ile Leu Leu Ala Ala Phe Gly Leu Leu Val
            85             90             95
Ala Leu Val Gly Ala Gln Cys Thr Asn Cys Val Gln Asp Asp Thr Ala
            100             105             110
Lys Ala Lys Ile Thr Ile Val Ala Gly Val Leu Phe Leu Leu Ala Ala
    115             120             125
Leu Leu Thr Leu Val Pro Val Ser Trp Ser Ala Asn Thr Ile Ile Arg
    130             135             140
Asp Phe Tyr Asn Pro Val Val Pro Glu Ala Gln Lys Arg Glu Met Gly
145             150             155             160
Ala Gly Leu Tyr Val Gly Trp Ala Ala Ala Ala Leu Gln Leu Leu Gly
            165             170             175
Gly Ala Leu Leu Cys Cys Ser Cys Pro Pro Arg Glu Lys Lys Tyr Thr
            180             185             190
Ala Thr Lys Val Val Tyr Ser Ala Pro Arg Ser Thr Gly Pro Gly Ala
    195             200             205
Ser Leu Gly Thr Gly Tyr Asp Arg Lys Asp Tyr Val
    210             215             220
```

What is claimed is:

1.   A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID

5   NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5, and a fragment of SEQ ID NO:7.

2.   A substantially purified variant having at least 90% amino acid identity to the amino acid sequence of claim 1.

10

3.   An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4.   An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

15

5.   An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6.   An isolated and purified polynucleotide having a sequence which is

20   complementary to the polynucleotide sequence of claim 3.

7.   An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, a fragment of SEQ ID NO:2, a fragment of SEQ ID NO:4, a fragment of SEQ ID

25   NO:6, and a fragment of SEQ ID NO:8.

8.   An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 7.

30   9.   An isolated and purified polynucleotide having a sequence which is

complementary to the polynucleotide of claim 7.

10.   An expression vector comprising at least a fragment of the polynucleotide of claim 3.

11.   A host cell comprising the expression vector of claim 10.

12.   A method for producing a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a fragment of SEQ ID NO:5, and a fragment of SEQ ID NO:7, the method comprising the steps of:

        a)   culturing the host cell of claim 11 under conditions suitable for the expression of the polypeptide; and

        b)   recovering the polypeptide from the host cell culture.

13.   A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

14.   A purified antibody which specifically binds to the polypeptide of claim 1.

15.   A purified agonist of the polypeptide of claim 1.

16.   A purified antagonist of the polypeptide of claim 1.

17.   A method for treating or preventing a disorder associated with increased or decreased apoptosis, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 13.

18.   A method for treating or preventing a disorder associated with increased or decreased apoptosis, the method comprising administering to a subject in need of such

treatment an effective amount of the antagonist of claim 16.

19.   A method for detecting a polynucleotide encoding the polypeptide comprising
the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3,
5   SEQ ID NO:5, SEQ ID NO:7, a fragment of SEQ ID NO:1, a fragment of SEQ ID NO:3, a
fragment of SEQ ID NO:5, and a fragment of SEQ ID NO:7 in a biological sample, the
method comprising the steps of:
            (a)   hybridizing the polynucleotide of claim 6 to at least one of the nucleic
acids in the biological sample, thereby forming a hybridization complex; and
10   (b)   detecting the hybridization complex, wherein the presence of the
hybridization complex correlates with the presence of the polynucleotide encoding the
polypeptide in the biological sample.

20.   The method of claim 19 wherein the nucleic acids of the biological sample are
15   amplified by the polymerase chain reaction prior to the hybridizing step.

ABSTRACT OF THE DISCLOSURE

The invention provides human apoptosis associated proteins (HAPOP) and

5   polynucleotides which identify and encode HAPOP. The invention also provides expression

vectors, host cells, antibodies, agonists, and antagonists. The invention also provides

methods for diagnosing, treating or preventing disorders associated with expression of

HAPOP.

FIGURE 1A

```
1    M E Y A M K S L S L L Y P K S L S R H V S V R T S V V T Q Q      642272
1    M D Y A M K S L S L L Y P R S L S R H V A S T A V V T Q Q        P56198
1    M E - - - - - - - - - - - V T G D A G V P E S G E I R T          GI 2065561

31   L S E P S P K A P R A R P C R V S T A D R S V R K G I M A        642272
31   L V S K P S R E T P R A R P C R V S T A D R K V R K G I M A      P56198
18   L - - - - - - - - - - K P C L L R R N Y S R E Q H G V A A        GI 2065561

61   Y S L E D L L K V R D T L M L A D K - - P F F L V L E E D        642272
61   H S L E D L L N K V Q D I L K L K D K - - P F S L V L E E D      P56198
37   S C L E D L R S K A C D I L A I D K S L T P V T L V L A E D      GI 2065561

89   G T V E T E E Y F Q A L A G D T V F M V L Q K G Q K W - -        642272
89   G T I V E T E E Y F Q A L A K D T M F M V L L K G Q K W - -      P56198
67   G T I V D D D D Y F L C L P S N T K F V A L A S N E K W A Y      GI 2065561

117  - - - - - - - - - Q P P S E Q G T R H P - L S L S H R P          642272
117  - - - - - - - - - K P P S E Q R K K R A Q L A L S Q K P          P56198
97   N N S D G G T A W I S Q E S F D V D E T D S G A G L K W K N      GI 2065561

135  A K - - - K I D V A R V T F D L Y K L N P Q D F I G - - C L      642272
136  T K - - - K I D V A R V T F D L Y K L N P Q D F I G - - C L      P56198
127  V A R Q L K E D L S S I I L - L S E E D L Q M L V D A P C S      GI 2065561
```

```
160  N V K A T F Y D T Y S - - - - L S Y D L H C C G A K R I - -          642272
161  N V K A T L Y D T Y S - - - - L S Y D L H C Y K A K R I - -          P56198
156  D L A Q E L R Q S C A T V Q R L Q H T L Q Q V L D Q R E E V          GI 2065561

184  - - M K E A F R W A L F S M Q A T G H V L L - - - - G T S C          642272
185  - - V K E I V R W T L F S M Q A T G H M L L - - - - G T S S          P56198
186  R Q S K Q L L Q L Y L Q A L E K E G S L L S K Q E E S K A A          GI 2065561

208  Y L Q - - - - - - - - - - - - - - - - - - Q L D A T                642272
209  Y M Q - - - - - - - - - - - - - - - - - - Q F L D A T              P56198
216  F G E E V D A V D T G I S R E T S S D V A L A S H I L T A L          GI 2065561

217  E G Q P P K - - - - - - - - - - - - - - - - - - - - -              642272
218  E E Q P A K - - - - - - - - - - - - - - - - - - - - -              P56198
246  R E K Q A P E L S L S S Q D L E L V T K E D P K A L A V A L          GI 2065561

224  - - - G K A S S L I P T C L K I L Q                                642272
225  - - A K P S S L L P A C L K M L Q                                  P56198
276  N W D I K K T E T V Q E A C E R E L A L R L Q Q T Q S L H S          GI 2065561

238                                                                    642272
239                                                                    P56198
306  L R S I S A S K A S P P G D L Q N P K R A R Q D P T                GI 2065561
```

FIGURE 1B

FIGURE 2A

```
1    M L Q E V S E Y L G V T S F R R K Y P D L E R R D L S H -      GI 1453807
1    M - - E Q C H N Y N A - - - - - - R L C A E R S V R - L P -    GI 606661

30   K E K L Y R E L N V I T E T Q C T L G T A - L R S D E          GI 1453807
21   - - - - F L D S Q T G V A Q S N C Y I W M E K R H R G P G      GI 606661

58   V I D L M I K E Y P A K H A E Y S V I L Q E K Q R I T          GI 1453807
46   L A S G Q L Y S Y P A R R - - - W R K R R A H P P              GI 606661

87   D H Y K - E Y S Q M Q Q Q N T Q K V E A S K V P E Y I K K      GI 1453807
69   E D P R L S F P S I K P D T D Q T L K K E G L - -              GI 606661

115  A A K A A E F N S L N R E R M E E R R A Y F D L Q T H          GI 1453807
92   I S Q D G S S L E A L L R T D P L E K R G A - - P D P R        GI 606661

144  V I Q V P Q G K Y K V L P T E R T K V S S Y P V A L I P G      GI 1453807
118  V D D D S L G E P P V - S N S R A R K R I - - - I E P D        GI 606661

173  Q F Q E Y Y K R Y S P D E L R Y L P L N T A L Y E P P L D      GI 1453807
142  D F - - - - - - - I E P                                        GI 606661

202  P E L P A L D S D G S D D G E D G R G D E K R N K G T          GI 1453807
144  - - L D D L D D E D Y E E D R P K R R G K S K S K G V          GI 606661
```

```
231  SDSSS---GNVSEGESPP---DSQEDSFQ          GI 1453807
171  SSARKKLDASILEDRDKPYACDICGKRYK          GI 606661

254  GR-------------QKSKDAAT-PRK            GI 1453807
200  NRPGLSYHYAHSHLAEEEGEDKEDSRPPT          GI 606661

268  DGPKRSVLSKSVPGYKPKVIPNAICGICL          GI 1453807
229  PVSQRSEEQKSKKGPDGLALPNNYCDFCL          GI 606661

297  -KGESNKKGKAESLHCSQCENSGHPSC            GI 1453807
258  GDSKINKKTGQPEELVSCSDCGRSGHPSC          GI 606661

325  LDMTMELVSMIKTYPWQCMECKTCIICGQ          GI 1453807
287  LQFTPVMMAAVKTYRWQCIECKCCNLCGT          GI 606661

354  PHHEEMMFCDMCDRGYHTFCV--GLGAI           GI 1453807
316  SENDQLLFCDDCDRGYHMYCLTPSMSEP           GI 606661

381  PSGRWICDCCQRAPPTPRKVGRRGKNSKE          GI 1453807
345  PEGSWSCHLCLDLL---KEKASIYQNQNS          GI 606661

410  G                                     1453807
371  S                                     GI 606661
```

FIGURE 2B

```
1   MANSGLQLLGFSMALLGWVGLVACTAIPQW    2059022
1   MSM-GLEITGTALAVLGWLGTIVCCALPMW    GI 2570129

31  QMSSYAGDNITAQAMYKGLWMDCVTQSTG     2059022
30  RVSAFIGSNIITSQNIWEGLWMNCVVQSTG    GI 2570129

61  MMSCKMYDSVLALSAALQATRALMVVSLVL    2059022
60  QMQCKVYDSLLALPQDLQAARALIVVAILL    GI 2570129

91  GFLAMFVATMGMKCTRCGGDDKVKARIAM     2059022
90  AAFGLLVALVGAQCTNCVQDDTAK-AKITI    GI 2570129

121 GGIFIVAGLAALVACSWYGHQIVTDFYN      2059022
119 VAGVLFLLAALTLVPVSWSANTIRDFYN      GI 2570129

151 PLIPTNIKYEFGPAIFIGWAGSALVILGGA    2059022
149 PVVPEAQKREMGAGLYVGWAAAALQLLGGA    GI 2570129

181 LLSCSCPGNE-----SKAGYRAPRSYPKSN    2059022
179 LLCCSCPPREKKYTATKVVYSAPRSTGPGA    GI 2570129

206 S------SKEYVV                     2059022
209 SLGTGYDRKDYVV                     GI 2570129
```

FIGURE 3

DECLARATION AND POWER OF ATTORNEY FOR
UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

HUMAN APOPTOSIS ASSOCIATED PROTEINS

the specification of which:

/___/ is attached hereto.

/ X / was filed on May 13, 1998 as application Serial No. 09/078,402 and if this box contains an X /___/, was amended on _____.

/___/ was filed as Patent Cooperation Treaty international application No. _____ on _____, 19__, if this box contains an X /_/, was amended on under Patent Cooperation Treaty Article 19 on _____19_, and if this box contains an X /_/, was amended on _____.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

1

Country	Number	Filing Date	Priority Claimed
			/_/ Yes /_/ No
			/_/ Yes /_/ No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)

I hereby appoint the following:

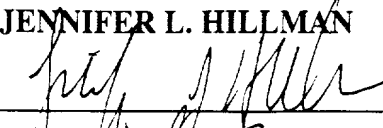**LUCY J. BILLINGS**   **Registration No. 36,749**
MICHAEL C. CERRONE   **Registration No. 39,132**
SHEELA MOHAN-PETERSON   **Registration No. 41,201**
COLETTE C. MUENZEN   **Registration No. 39,784**
KAREN J. ZELLER   **Registration No. 37,071**

respectively and individually, as my attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

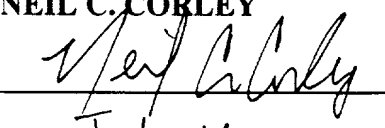<div align="center">

SHEELA MOHAN-PETERSON, ESQ.
INCYTE PHARMACEUTICALS, INC.
3174 PORTER DRIVE, PALO ALTO, CA 94304
TEL: 650-855-0555   FAX: 650-845-4166

</div>

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.
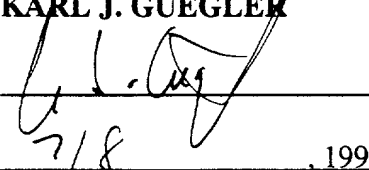
Sole Inventor or
First Joint Inventor:

Full name: **JENNIFER L. HILLMAN**

Signature: _____

Date: _____ , 1998

Citizenship: United States of America

Residence: Mountain View, California

P.O. Address: 230 Monroe Drive, #12
Mountain View, California

Second Joint Inventor:

Full name: **NEIL C. CORLEY**

Signature: _____

Date: July 14 , 1998

Citizenship: United States of America

Residence: Mountain View, California

P.O. Address: 1240 Dale Avenue, #30
Mountain View, California, 94040

Third Joint Inventor:

Full name: **KARL J. GUEGLER**

Signature: _____

Date: 7/8 , 1998

Citizenship: Switzerland

Residence: Menlo Park, California

P.O. Address: 1048 Oakland Avenue
Menlo Park, California, 94025

3

Fourth Joint Inventor:

Full name: **CHANDRA PATTERSON**

Signature: *Chandra Patterson*

Date: *July 7* , 1998

Citizenship: United States of America

Residence: Mountain View, California

P.O. Address: 2189 Leland Avenue
Mountain View, California, 94040

Fifth Joint Inventor:

Full name: **MARIAH R. BAUGHN**

Signature: *Mariah R Baughn*

Date: *July 7* , 1998

Citizenship: United States of America

Residence: San Leandro, California

P.O. Address: 14244 Santiago Road
San Leandro, California, 94577

4

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Hillman et al.

Title: HUMAN APOPTOSIS ASSOCIATED PROTEINS

Serial No.: 09/078,402   Filing Date: May 13, 1998

Examiner: Worral, T.   Group Art Unit: 1642

Assistant Commissioner for Patents
Washington, D.C. 20231

CERTIFICATE UNDER 37 C.F.R. §3.73(b), REVOCATION OF POWER OF ATTORNEY AND APPOINTMENT OF NEW ATTORNEYS

Sir:

The undersigned has reviewed all the documents in the chain of title of the above-identified patent application and, to the best of undersigned's knowledge and belief, title is in the assignee identified above.

Incyte Pharmaceuticals, Inc., having a principal place of business located at 3174 Porter Drive, Palo Alto, California 94304, certifies that it is the assignee and owner of the entire right, title and interest in, to, and under the invention described and claimed in the above-identified application by virtue of an Assignment recorded at Reel 9336, Frame 0300, hereby revokes all previous powers of attorney and appoints the following patent attorneys/agents:

Narinder S. Banait	**Reg. No. 43,482**
Adam Warwick Bell	**Reg. No. 43,490**
Lucy J. Billings	**Reg. No. 36,749**
Michael C. Cerrone	**Reg. No. 39,132**
Diana Hamlet-Cox	**Reg. No. 33,302**
Colette C. Muenzen	**Reg. No. 39,784**
Lynn E. Murry	**Reg. No. 42,918**

48204   1   09/078,402

Danielle M. Pasqualone   Reg. No. 43,847
Susan K. Sather   Reg. No. 44,316
David G. Streeter   Reg. No. 43,168

Please direct all correspondence to:

<div align="center">

Legal Department
Incyte Pharmaceuticals, Inc.
3174 Porter Drive
Palo Alto, California 94304

</div>

and direct all telephone calls and facsimile transmissions to: <u>Diana Hamlet-Cox, Ph.D.</u>,
Incyte Pharmaceuticals, Inc., Phone: (650) <u>845-4639</u> Fax: (650) 849-8886.

The undersigned (whose title is supplied below) is empowered to act on behalf of the assignee.

I hereby declare that all statements made herein of my own knowledge are true, and that these statements are made with the knowledge that willful false statements, and the like so made, are punishable by fine or imprisonment, or both, under Section 1001, Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

<div align="center">

INCYTE PHARMACEUTICALS, INC.

</div>

Date: <u>October 11, 1999</u>   By: _____
                                      Lee Bendekgey
                                      VP, General Counsel/Corporate Secretary

09/078,402